Polarimetric and birefringence analysis of presumed amyloid-β deposits in the retina in association with Alzheimer's disease

by

Tao Jin

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Physics (Nanotechnology)

Waterloo, Ontario, Canada, 2018

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

The thesis author is the sole author of all written chapters. Coauthors for earlier materials presented at conferences:

Coauthors for Frontiers in Optics 2016 on which I was the first author (*Mapping the birefringence of amyloid deposits found in retinas in association with Alzheimers disease*) are Laura Emptage, David DeVries, and Melanie CW Campbell.

Coauthors for Annual Meeting of the Association for Research in Vision and Ophthalmology (ARVO 2017) on which I was the first author (*The effects of amyloid-β deposits on retinal tissue assessed from polarization properties*) are Laura Emptage, David DeVries, Monika Kitor, and Melanie CW Campbell.

The retinal samples were prepared by Monika Kitor and Laura Emptage. The pure amyloid samples were prepared by Frank Corapi and Laura Emptage. The Mueller matrix microscope used to acquire the 16 intensity images was automated by Frank Corapi and Riley Delaney. The code for image registration of the acquired 16 images from the Mueller matrix microscope was written by Erik Mason. The images used for testing different segmentation algorithms were from a database which was created and organized by Erik Mason. I would like to acknowledge helpful discussion with David DeVries, Erik Mason, Laura Emptage, Rachel Redekop, Alexander Hannides Rajapakse, Ji Ren, Dr.Campbell. Dr.Campbell made helpful comments and adjustments to the thesis. Erik Mason, Laura Emptage, Monika Kitor, Alexander Hannides Rajapakse made editorial changes to the text of the thesis.

This research is supported by Natural Sciences and Engineering Research Council (NSERC), Canadian Institutes of Health Research (CIHR), and the University of Waterloo.

Abstract

Alzheimer's disease is a neurodegenerative disease which leads to symptoms such as loss of memory and other neurological dysfunctions, ultimately leading to death. Two key hallmarks of Alzheimer's disease are abnormally folded Amyloid-β and tau proteins in the brain. Currently, the diagnosis of Alzheimer's disease is only confirmed by finding these hallmarks in the brain after death. Two of the methods that reach a probable diagnosis of Alzheimer's disease are positron emission tomography and cerebral spinal fluid analyses. However, both methods are invasive as positron emission tomography uses radiation and dyes, and examination of cerebral spinal fluid requires invasive extraction from the patients' body. Fortunately, amyloid deposits, which presumed to contain Amyloid-β, have been found in the retina. *In vivo* imaging of amyloid deposits using curcumin staining has been shown in AD patients receiving oral curcumin for several days. Since the retina can be directly imaged through the pupil and polarimetry does not require a dye, a noninvasive method of diagnosis could combine polarimetry and a retinal imaging device.

In this thesis, one of the key polarimetric properties of amyloid deposits, linear birefringence, was studied. The linear birefringence was computed by using the combination of Mueller Matrix Polarimetry for linear retardance and Confocal Laser Scanning Microscopy for thickness. These measurements were conducted in retinal deposits as well as a matched number of pure Amyloid-β deposits with similar thickness grown on glass slides. Between the two types of deposits, there was no significant difference shown in linear birefringence fitted to the linear regions of the retardance versus thickness plot. Both types of amyloid deposits show similar trends at high thickness where linear retardance plateaus and then decreases with increasing thickness, giving low birefringence value at high thickness. This result suggests that the retinal deposits are composed primarily of amyloid-β and that there is only short range order of the fibrils in the retinal deposits. The birefringence of retinal deposits is also higher than the background retinal nerve fiber layer and the senile plaques in the brain measured to date. These senile plaques are higher in thickness than the retinal deposits and have birefringence values similar to the low birefringence values of pure Amyloid-β deposits of high thickness. The high birefringence of the retinal deposits demonstrates the feasibility of *in vivo* imaging of retinal amyloid deposits using a patented label free method.

In order to calculate the retardance used here to infer the birefringence measurements and in future clinical applications, the Mueller matrix computed from the acquired images needs to be decomposed. However, for a large image, the decomposition may take a long time. Hence, an accelerated implementation of the polar decomposition was developed and used in this thesis. By implementing the accelerated decomposition on a graphic processing unit, the speed of decomposition is more than 15 times faster than the original calculation

before acceleration.
For computing the birefringence of amyloid deposits, image segmentation was conducted to separate them from the surrounding retina. Automatic segmentation methods including Otsu's method, K-means clustering, Gaussian mixture model, and Markov random field were tested. Gaussian mixture model and Markov random field showed the highest success rate while both Otsu's method and K-means clustering performed poorly. It is thus recommended to use either Gaussian mixture model (faster) or Markov random field (slightly more accurate) to segment images of amyloid deposits.
The results reported in this thesis show polarization properties of amyloid deposits, such as strong birefringence values, which are useful in identifying these deposits in the retina. Coupled with fast decomposition and automatic segmentation of the deposits in the image, an efficient, noninvasive diagnostic method may be established in living eyes in the future.

Acknowledgements

I would like to thank my parents for supporting my study overseas. I would like to thank Dr. Campbell for supervising me for my research in Campbell's lab. I would like to thank group members in the lab who have offered tremendous help during my time in the lab.

Dedication

Dedicated to my parents and my friends

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AD	Alzheimer's Disease
PD	Parkinson's Disease
DLB	Dementia with Lewy Bodies
APOE	Apolipoprotein E
Aβ	Amyloid-beta
NFTs	Neurofibrillary Tangles
PET	Positron Emission Tomography
CSF	Cerebral Spinal Fluid
RNFL	Retinal Nerve Fiber Layer
AFM	Atomic Force Microscopy
MMP	Mueller Matrix Polarimetry
PSG	Polarization State Generator
PSA	Polarization State Analyzer
QWP	Quarter Wave Plate
OCT	Optical Coherence Tomography
PSOCT	Polarization Sensitive Optical Coherence Tomography
PD	Polar decomposition
CLSM	Confocal Laser Scanning Microscope
GPU	Graphical Processing Unit
CUDA	Compute Unified Device Architecture
SM	Streaming Multiprocessor
SP	Streaming Processor
GMM	Gaussian Mixture Model
MRF	Markov Random Field
ML	Maximum Likelihood
EM	Expectation Maximization
ICM	Iterative Conditional Mode
HMCM	Hidden Markov Chain Model

HHMM	Hidden Hierarchical Markov Model
CNN	Convolutional Neural Network
SVM	Support Vector Machine

Chapter 1

Introduction

1.1 Introduction of AD

Alzheimer's disease (AD) is a fatal neural degenerative disease, which is also the most prevalent cause of dementia [1]. Dementia is manifested in a group of symptoms such as difficulties in memory, language, and other cognitive impairments [2]. Other diseases also causing dementia include vascular dementia, Dementia with Lewy bodies (DLB), Parkinson's disease (PD), etc., each having different types of symptoms. Specifically for AD patients, early signs include difficulties in remembering recent events, depression, and apathy. With the progression of AD, more severe symptoms, such as impairment in communication, difficulties in speaking, swallowing and walking, etc., may show up. The percentage of population above 60 years old affected by AD is 5% to 7% [3]. The global population impacted by AD was about 35.6 million in 2010, and estimated to triple by 2050 [4]. Risk factors of AD can be categorized as age, genetic, and environmentally related [5]. The increased percentage of population with dementia at a higher age indicates that age is a strong risk factor [4]. Besides age, carriers of Apolipoprotein E4 (APOE4) gene, which disrupts Amyloid-β (Aβ) clearance from the brain, are several times more likely to develop AD than noncarriers [6]. In addition, environmental factors, such as physical inactivity, lack of social engagement, obesity, diabetes, and hypertension, also add to the risk of developing AD.

AD is characterized by two pathological hallmarks: misfolded Aβ protein in amyloid plaques and tau protein in neurofibrillary tangles (NFTs). Currently, AD diagnosis is only confirmed by finding these two hallmarks in the patient's brain after death. Aβ is considered as the driver of disease progression. This is known as the amyloid cascade hy-

pothesis [7]. Evidence has suggested that the soluble Aβ 42 oligomers can cause damage to the lipid membranes [8]. These oligomers later assemble into fibrils which then form senile plaques containing these Aβ fibrils. The amyloid cascade hypothesis [7] also suggests that tau is a downstream mechanism to Aβ while others suggest that tau acts in parallel with Aβ which amplifies its neurotoxicity [9].
The current probable diagnosis of AD is done with Positron Emission Tomography (PET) [10] or examination of cerebral spinal fluid (CSF) [11]. However, PET scans are expensive and use radiation. Analyzing CSF requires invasive extraction of fluid surrounding the patients' spine. These two diagnostic methods also usually take place after the patient shows symptoms of cognitive decline. Currently, no effective treatment has been established. Some suggest the reason for this lack of successful treatment is that presently diagnosis does not take place early enough to influence its progression [12]. The two hallmarks of AD, Aβ and tau, accumulate in the brain before the onset of the symptoms of cognitive impairment. This symptomless period is called prodromal AD. Therefore, it is important to establish the methods for early stage diagnosis so that treatments may focus on blocking the pathway of Aβ and tau before their neurotoxicity degrades the patient's cognitive functions.

1.2 Amyloid beta in the retina and its birefringence properties

1.2.1 Birefringence of retinal amyloid deposits

An early report [13] has observed evidence of ocular abnormalities in AD patients. Since retinal tissue includes neural tissue as an outgrowth from the central nervous system, its structural change in AD has been extensively investigated. These changes include thinning of the retinal nerve fiber layer (RNFL) [14], reduction of the number of retinal ganglion cells [15], and vascular changes (decreased blood flow, reduced vessel thickness) [14, 16]. In addition to changes in retinal structure, amyloid deposits, which presumed to contain Aβ fibrils, have been found in the retina of humans [17, 18] while others do not find them [19]. The morphology of amyloid deposits has been shown using Atomic Force Microscopy (AFM) and fluorescence imaging [17]. Koronyo *et. al.* [20] have shown *in vivo* detection of the amyloid deposits in human retina using curcumin staining. However, patients had to ingest curcumin for 2 to 10 days to achieve binding of the deposits in retina.
Further, the amyloid deposits have shown birefringent properties under Congo Red staining [21, 22, 23]. Birefringence is the property of the material having different refractive

index along orthogonal direction. The birefringence of senile plaques related to AD and other neurodegenerative diseases stained with Congo Red has been shown [24]. The senile plaques of AD shows radial symmetry in their fast axis and transmission axis (introduced in Section 1.3.2). Since the amyloid plaques are composed of protein fibrils, it was postulated by Campbell [25] that it might be naturally birefringent under polarized light. In fact, the natural birefringence as well as other polarization properties of amyloid deposits from post mortem retina have been shown in animal model [26] and human [27, 28]. Recently, the birefringence of unstained senile plaques has also been imaged by polarization sensitive optical coherence tomography (PSOCT) [29]. Hence, the amyloid fibrils, when forming plaques, can naturally have birefringent properties which can be used for label free detection in the retina.
The birefringence property can be interpreted as intrinsic birefringence and form birefringence. Intrinsic birefringence is a property independent of the refractive index of the medium [30]. It is caused by anisotropic arrangement of the atoms in the crystal. Form birefringence, on the other hand, results from ordered alignment of macromolecules or fibrils in the environment. The difference between the refractive index of the fibrils and the environment gives a different refractive index in parallel and perpendicular directions to the fibril axis. Since the birefringence of amyloid deposits may results from both the intrinsic and form birefringence, the theory of these two types of birefringence are introduced here.

1.2.2 Theory of intrinsic birefringence

In electrodynamics, the electric displacement field $\boldsymbol{D}$ is related to the electric field $\boldsymbol{E}$ as

$$\boldsymbol{D} = \varepsilon_0 \varepsilon \boldsymbol{E} \tag{1.1}$$

where ε_0 is the vacuum permitivity and ε is the relative permitivity. If the arrangement of atoms in the crystal is isotropic, the ε can be written as a scalar times an identity matrix. A representative isotropic crystal structure is the cubic crystal. However, if the arrangement of atoms is anisotropic, ε becomes a symmetric dielectric tensor $\boldsymbol{\varepsilon}$ [31]. Thus, Eq.1.1 becomes

$$\begin{pmatrix} D_x \\ D_y \\ D_z \end{pmatrix} = \varepsilon_0 \begin{pmatrix} \varepsilon_{xx} & \varepsilon_{xy} & \varepsilon_{xz} \\ \varepsilon_{yx} & \varepsilon_{yy} & \varepsilon_{yz} \\ \varepsilon_{zx} & \varepsilon_{zy} & \varepsilon_{zz} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} \tag{1.2}$$

where the off diagonal entries have the relationship

$$\varepsilon_{ij} = \varepsilon_{ji} \tag{1.3}$$

where $i \neq j \in \{x, y, z\}$. This matrix $\boldsymbol{\varepsilon}$ can be diagonalized using eigenvalue decomposition. Its diagonlized matrix $\boldsymbol{\varepsilon}'$ is

$$\boldsymbol{\varepsilon}' = \begin{pmatrix} \varepsilon_1 & 0 & 0 \\ 0 & \varepsilon_2 & \\ 0 & 0 & \varepsilon_3 \end{pmatrix} \tag{1.4}$$

If $\varepsilon_1 = \varepsilon_2 = \varepsilon_3$, the crystal is isotropic. However, if only two values in $\{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$ are equivalent, the crystal becomes a uniaxial crystal. For example, if $\varepsilon_1 = \varepsilon_2 \neq \varepsilon_3$, then the direction along the corresponding eigenvector of ε_3 is the optic axis of the crystal. If the polarization of the input light is parallel to this axis, the light experiences one refractive index $n_e = \sqrt{\varepsilon_3}$. If the polarization of the input light is perpendicular to this axis, it will experience another refractive index $n_o = \sqrt{\varepsilon_1} = \sqrt{\varepsilon_2}$. Then the intrinsic birefringence Δn of this uniaxial crystal is

$$\Delta n = n_e - n_o \tag{1.5}$$

Uniaxial crystal systems include trigonal, tetragonal, and hexagonal [31]. Further, if all three eigenvalues are different, the crystal is a biaxial crystal.

1.2.3 Theory of form birefringence

Form birefringence results from the ordered alignment of molecules or fibrils with similar dielectric constant inside an environment with another dielectric constant. As shown in Fig.1.1, the fibrils have a dielectric constant of ϵ_1 and width of t_1 and the spacing between fibrils has a dielectric constant of ϵ_2 and a width of t_2. The form birefringence is then derived using the boundary condition in electrodynamics [30].

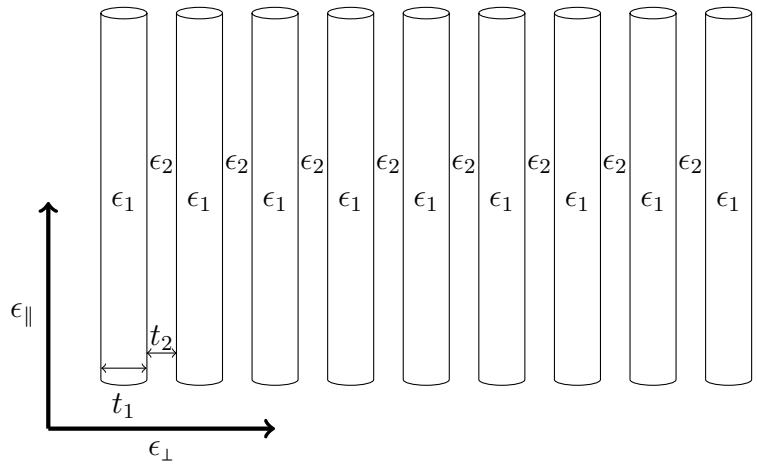


Figure 1.1: Ordered alignment of fibrils with dielectric constant ϵ_1 inside an environment with a dielectric constant of ϵ_2. The thickness of the fibrils is t_1 and the spacing between them is t_2. $\epsilon_\parallel$ and $\epsilon_\perp$ are the dielectric constant parallel and perpendicular to the axis of the fibrils.

First, assuming there is no surface charge, the normal component of the electric displacement $\boldsymbol{D}_\perp$ is continuous across the boundary. So the normal component of electric field in the fibrils $\boldsymbol{E}_{\perp_1}$ and the electric field in the spacing between the fibrils $\boldsymbol{E}_{\perp_2}$ are written as

$$\boldsymbol{E}_{\perp_1} = \frac{\boldsymbol{D}_\perp}{\epsilon_1} \qquad \boldsymbol{E}_{\perp 2} = \frac{\boldsymbol{D}_\perp}{\epsilon_2} \tag{1.6}$$

So the mean electric field normal to the axis of the fibrils in Fig.1.1 can be computed as

$$\begin{aligned} \boldsymbol{E}_\perp =& \frac{t_1 \boldsymbol{E}_{\perp_1} + t_2 \boldsymbol{E}_{\perp_2}}{t_1 + t_2} \\ =& \frac{t_1}{t_1 + t_2} \frac{\boldsymbol{D}_\perp}{\epsilon_1} + \frac{t_2}{t_1 + t_2} \frac{\boldsymbol{D}_\perp}{\epsilon_2} \end{aligned} \tag{1.7}$$

Hence, the normal dielectric constant $\epsilon_\perp$ is computed using $\boldsymbol{D}_\perp$ and $\boldsymbol{E}_\perp$ as

$$\begin{aligned} \epsilon_\perp =& \frac{\boldsymbol{D}_\perp}{\boldsymbol{E}_\perp} \\ =& \frac{\epsilon_1 \epsilon_2}{\frac{t_1}{(t_1+t_2)}\epsilon_1 + \frac{t_2}{(t_1+t_2)}\epsilon_2} \\ =& \frac{\epsilon_1 \epsilon_2}{f_1 \epsilon_2 + f_2 \epsilon_1} \end{aligned} \tag{1.8}$$

where f_1 and f_2 are the volume fraction of the fibrils and the spacing between fibrils written as

$$f_1 = \frac{t_1}{(t_1 + t_2)} \qquad f_2 = \frac{t_2}{(t_1 + t_2)} \tag{1.9}$$

According to the boundary condition in electrodynamics, the parallel component of the $\boldsymbol{E}_\parallel$ is continuous across the boundary. So the parallel component of the displacement field $\boldsymbol{D}_{\parallel_1}$ in fibrils and $\boldsymbol{D}_{\parallel_2}$ are

$$\boldsymbol{D}_{\parallel_1} = \epsilon_1 \boldsymbol{E}_\parallel \qquad \boldsymbol{D}_{\parallel_2} = \epsilon_2 \boldsymbol{E}_\parallel \tag{1.10}$$

Then the mean displacement field parallel to the axis of fibrils in Fig.1.1 is computed as

$$\begin{aligned} \boldsymbol{D}_\parallel =& \frac{t_1 \boldsymbol{D}_{\parallel_1} + t_2 \boldsymbol{D}_{\parallel_2}}{t_1 + t_2} \\ =& \frac{t_1 \epsilon_1 \boldsymbol{E}_\parallel + t_2 \epsilon_2 \boldsymbol{E}_\parallel}{y} \\ =& f_1 \epsilon_1 \boldsymbol{E}_\parallel + f_2 \epsilon_2 \boldsymbol{E}_\parallel \end{aligned} \tag{1.11}$$

Thus, the dielectric constant parallel to the fibril axis can be written as

$$\begin{aligned} \epsilon_\parallel =& \frac{\boldsymbol{D}_\parallel}{\boldsymbol{E}_\parallel} \\ =& f_1 \epsilon_1 + f_2 \epsilon_2 \end{aligned} \tag{1.12}$$

Hence, the difference between the parallel and perpendicular components of the dielectric constant $\Delta\epsilon$ is

$$\begin{aligned} \Delta\epsilon =& \epsilon_\parallel - \epsilon_\perp \\ =& \frac{f_1 f_2 (\epsilon_1 - \epsilon_2)^2}{f_1 \epsilon_2 + f_2 \epsilon_1} \geqslant 0 \end{aligned} \tag{1.13}$$

Since the dielectric constant is the square of refractive index, then

$$\epsilon_1 = n_1^2 \qquad \epsilon_2 = n_2^2 \tag{1.14}$$

and

$$\epsilon_\parallel = n_\parallel^2 \qquad \epsilon_\perp = n_\perp^2 \tag{1.15}$$

So Eq.1.13 can be expressed as

$$n_\parallel^2 - n_\perp^2 = \frac{f_1 f_2 (n_1^2 - n_2^2)^2}{f_1 n_2^2 + f_2 n_1^2} \geqslant 0 \tag{1.16}$$

So the refractive index along the direction parallel to the axis of the fibrils will be larger than the perpendicular direction. Then the slow axis of this assembly of fibrils will be aligned with the axis of the fibrils. Hence, the form birefringence Δn can be expressed as

$$\begin{aligned}\Delta n =& n_{\parallel} - n_{\perp} \\ =& \sqrt{f_1 n_1^2 + f_2 n_2^2} - \sqrt{\frac{n_1^2 n_2^2}{f_1 n_2^2 + f_2 n_1^2}}\end{aligned} \tag{1.17}$$

Additionally, if there are multiple layers of these ordered alignments along the depth direction in Fig.1.1, the birefringence value is adjusted to $\Delta n'$ by multiplying a filling factor $h_l/(h_l + h_s)$ as [32]

$$\Delta n' = \frac{h_l}{h_l + h_s}\Delta n \tag{1.18}$$

where h_l is the thickness of the layer and h_s is the spacing between the layers. Hence, if the fibrils are more closely packed along the depth direction, the measured birefringence can increase.

1.3 Polarized light and Mueller-Stokes formalism

The birefringence property of amyloid enables polarized light to be a tool for detection. The birefringence imposes phase retardation between field components of the input polarized light parallel to the extraordinary axis and the ordinary axis of the medium. Hence, using polarized light to interact with the sample, one can probe its birefringence and other anisotropic properties. To represent the polarized light and its interaction with the sample, two mathematical tools are available: the Jones matrix formalism and the Mueller-Stokes formalism. The instrument used in this thesis is based on the Mueller-Stokes formalism. Although the Mueller-Stokes formalism requires more images to be taken than the Jones matrix formalism, it can compute the depolarization effect of the medium imposed on the input polarized light, which Jones matrix formalism is not able to do.
The introduction of polarized light and derivation of the Mueller-Stokes formalism given in this subsection is constrained in the following 2 conditions:

1. The light is quasi monochromatic which means the spectral width $\Delta\nu$ is much lower than the mean frequency $\bar{\nu}$. Hence, the coherence time $1/\Delta\nu$ is much longer than the period $1/\bar{\nu}$ of the wave. The coherence time is the duration in which the polarization ellipse (introduced in this subsection) is stable.

2. The polarization is 2D which means the plane containing the polarization ellipse does not change over time. For 3D polarization, one can refer to Chapter 2 in Reference [33].

1.3.1 Polarized light and Stokes vector

First, for a plane wave propagating in the direction of z axis, its component in x axis $E_x(t)$ and y axis $E_y(t)$ are

$$\begin{aligned} E_x(t) &= A_x(t)\cos[\tau(t)+\delta_x(t)] \\ E_y(t) &= A_y(t)\cos[\tau(t)+\delta_y(t)] \end{aligned} \tag{1.19}$$

where $A_x(t)$, $A_y(t)$ are the amplitude of the wave in x and y directions and $\delta_x(t)$ and $\delta_y(t)$ are the phase factors of the wave. $\tau(t)$ is the propagator term, which in the case of quasi monochromatic light, can be expressed as

$$\tau(t) = \bar{k}z - \bar{\omega}t \tag{1.20}$$

where $\bar{k}$ and $\bar{\omega}$ are the mean wave vector and mean angular frequency.
Dividing by amplitudes in Eq.1.19, one gets

$$\begin{aligned} \frac{E_x(t)}{A_x(t)} &= \cos\tau(t)\cos\delta_x(t) - \sin\tau(t)\sin\delta_x(t) \\ \frac{E_y(t)}{A_y(t)} &= \cos\tau(t)\cos\delta_y(t) - \sin\tau(t)\sin\delta_y(t) \end{aligned} \tag{1.21}$$

Then by combining $\delta_x(t)$ and $\delta_y(t)$, one gets

$$\begin{aligned} \frac{E_x(t)}{A_x(t)}\sin\delta_y(t) - \frac{E_y(t)}{A_y(t)}\sin\delta_x(t) &= \cos\tau(t)\sin(\delta_y(t)-\delta_x(t)) \\ \frac{E_x(t)}{A_x(t)}\cos\delta_y(t) - \frac{E_y(t)}{A_y(t)}\cos\delta_x(t) &= \sin\tau(t)\sin(\delta_y(t)-\delta_x(t)) \end{aligned} \tag{1.22}$$

Squaring and adding the term in Eq.1.22 to get rid of the term τ gives

$$\left(\frac{E_x(t)}{A_x(t)}\right)^2 + \left(\frac{E_y(t)}{A_y(t)}\right)^2 - 2\frac{E_x(t)E_y(t)}{A_x(t)A_y(t)}\cos\delta(t) = (\sin\delta(t))^2 \tag{1.23}$$

where $\delta(t) = \delta_y(t) - \delta_x(t)$. It is clear that this equation describes a complete ellipse if the terms $A_x(t)$, $A_y(t)$ and $\delta(t)$ are constant in a period of the wave. This ellipse, shown in

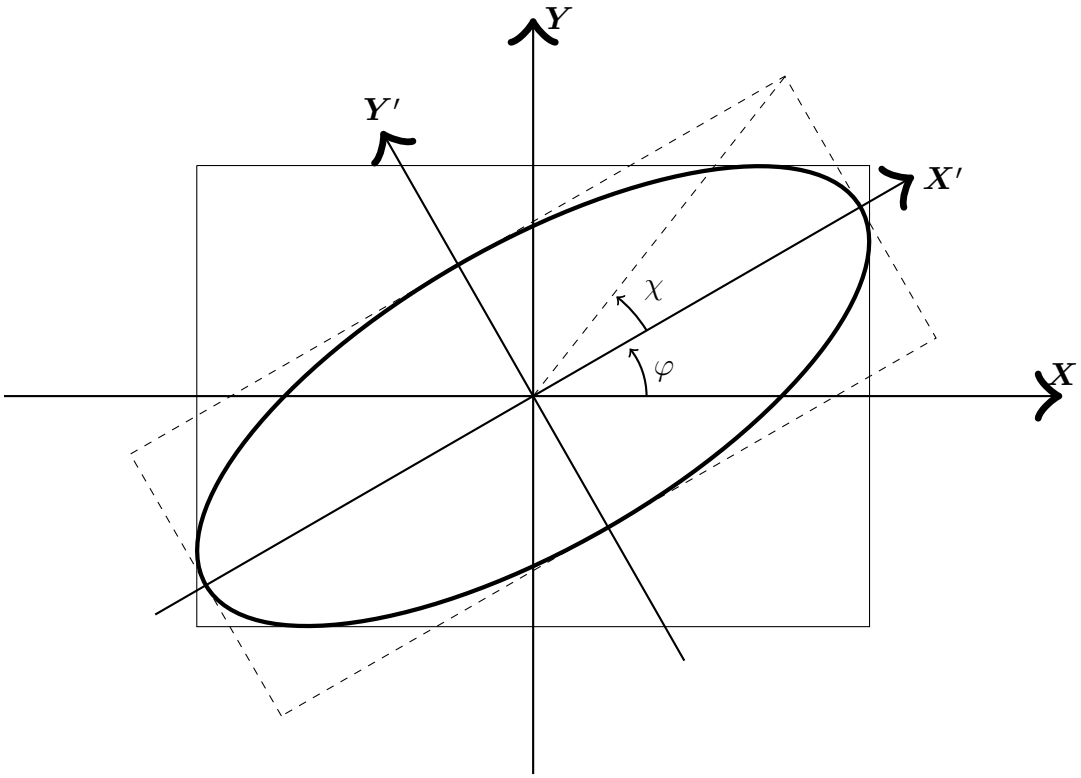


Figure 1.2: Polarization ellipse described by Eq.1.23 with azimuth and ellipticity computed as Eq.1.24,1.25

Fig.1.2, is called polarization ellipse which is the shape of the evolution of the endpoint of the vector $(E_x(t), E_y(t))$ in a period. Azimuth and ellipticity are usually used to describe the orientation and shape of an ellipse. Here, its azimuth $\psi(t)$ and ellipticity $\chi(t)$ are computed as [30]

$$\psi(t) = \frac{1}{2}\arctan\Big[\frac{2A_x(t)A_y(t)}{A_x(t)^2 - A_y(t)^2}\cos\delta(t)\Big] \tag{1.24}$$

$$\chi(t) = \frac{1}{2}\arcsin\Big[\frac{2A_x(t)A_y(t)}{A_x(t)^2 + A_y(t)^2}\sin\delta(t)\Big] \tag{1.25}$$

For a monochromatic light in vacuum, the amplitudes $A_x(t)$, $A_y(t)$ and the phase $\delta(t)$ stay constant. Hence, the shape and size of the ellipse is constant throughout its propagation. So, monochromatic light is always completely polarized. However, for quasi monochromatic light, the amplitudes and phase can fluctuate with time. Thus, over the measurement time

T, which is usually longer than the coherence time of quasi monochromatic light [34], the shape of the ellipse could fluctuate. Only if the ratio of $A_y(t)/A_x(t)$ and the phase $\delta(t)$ remain unchanged will the quasi monochromatic light be completely polarized.
Continuing with Eq.1.23, multiplying both sides by $4A_x^2(t)A_y^2(t)$ and taking the time average yields

$$\begin{aligned} &4\langle A_y^2(t)E_x^2(t)\rangle + 4\langle A_x^2(t)E_y^2(t)\rangle - 8\langle A_x(t)A_y(t)E_x(t)E_y(t)\cos\delta(t)\rangle \\ =&4\langle (A_x(t)A_y(t)\sin\delta(t))^2\rangle \end{aligned} \tag{1.26}$$

The time average of a term $f(t)$ is computed as

$$\langle f(t)\rangle = \lim_{T\to\infty}\frac{1}{T}\int_0^T f(t)dt \tag{1.27}$$

Hence, if the light is monochromatic, each term in Eq.1.26 can be computed by combining with Eq.1.19,

$$\langle A_y^2(t)E_x^2(t)\rangle = \frac{1}{2}A_x^2A_y^2 \tag{1.28}$$

$$\langle A_x^2(t)E_y^2(t)\rangle = \frac{1}{2}A_x^2A_y^2 \tag{1.29}$$

$$\langle A_x(t)A_y(t)E_x(t)E_y(t)\cos\delta(t)\rangle = \frac{1}{2}A_x^2A_y^2\cos^2\delta \tag{1.30}$$

Plugging these terms back to Eq.1.26 gets

$$4A_x^2A_y^2 - 4(A_xA_y\cos\delta)^2 = (A_xA_y\sin\delta)^2 \tag{1.31}$$

which is equivalent to

$$(A_x^2+A_y^2)^2 - (A_x^2-A_y^2)^2 - (2A_xA_y\cos\delta)^2 = (2A_xA_y\sin\delta)^2 \tag{1.32}$$

Then, the Stokes vector for a monochromatic light is defined as each term in the brackets in Eq.1.32 as

$$\begin{aligned} S_0 &= A_x^2 + A_y^2 \\ S_1 &= A_x^2 - A_y^2 \\ S_2 &= 2A_xA_y\cos\delta \\ S_3 &= 2A_xA_y\sin\delta \end{aligned} \tag{1.33}$$

It is clear that for monochromatic light,

$$S_0^2 - S_1^2 - S_2^2 - S_3^2 = 0 \tag{1.34}$$

The light is completely polarized. For a quasi monochromatic light, the general formalism of Stokes vector in Eq.1.33 is written as measurement of each term

$$\begin{aligned} S_0 &= \langle A_x(t)^2 \rangle + \langle A_y(t)^2 \rangle \\ S_1 &= \langle A_x(t)^2 \rangle - \langle A_y(t)^2 \rangle \\ S_2 &= 2\langle A_x(t) A_y(t) \cos\delta(t) \rangle \\ S_3 &= 2\langle A_x(t) A_y(t) \sin\delta(t) \rangle \end{aligned} \tag{1.35}$$

which can also be expressed as [30]

$$\begin{aligned} S_0 &= \langle I_x \rangle + \langle I_y \rangle \\ S_1 &= \langle I_x \rangle - \langle I_y \rangle \\ S_2 &= \langle I_{+45^\circ} \rangle - \langle I_{-45^\circ} \rangle \\ S_3 &= \langle I_R \rangle - \langle I_L \rangle \end{aligned} \tag{1.36}$$

where the I_x, I_y, I_{+45° and I_{-45° are the intensity of light along x axis, y axis, +45° axis and −45° axis. I_R and I_L are the intensity of light of right circularly polarized light and left circularly polarized light. Hence, the 2nd to 4th entry of the Stokes vector can be deemed as the difference in the intensity of each pair of orthogonal polarization.
For Eq.1.35 to satisfy Eq.1.33, one can arrive at the same condition as Eq.1.24 and 1.25, that is the ratio of $A_y(t)/A_x(t)$ and the phase $\delta(t)$ has to be constant for the quasi monochromatic light to be completely polarized. If Eq.1.33 is not satisfied, the light is partially polarized. The degree of polarization of this partially polarized light is defined as

$$P = \frac{\sqrt{S_1^2 + S_2^2 + S_3^2}}{S_0} \tag{1.37}$$

Although the shape of the polarization ellipse fluctuates, partial polarized light maintains its degree of polarization during measurement.
Using Eq.1.24,1.25, the Stokes vector in Eq.1.35 can be written as

$$\boldsymbol{S} = S_0 \begin{pmatrix} 1 \\ P\cos 2\psi \cos 2\chi \\ P\sin 2\psi \cos 2\chi \\ P\sin 2\chi \end{pmatrix} \tag{1.38}$$

where P, ψ and χ are the degree of polarization, azimuth, and ellipticity of the polarized light. The elements in Eq.1.38 can form the vector $(1, P\boldsymbol{s})^T$ where $\boldsymbol{s}$ is

$$\boldsymbol{s} = (\cos 2\psi \cos 2\chi, \sin 2\psi \cos 2\chi, \sin 2\chi) \tag{1.39}$$

$\boldsymbol{s}$ can be represented by a point on a sphere. This sphere is called the Poincaré sphere which describes every possible state of polarization of a completely polarized light [30]. The Poincaré sphere is drawn in Fig.1.3.

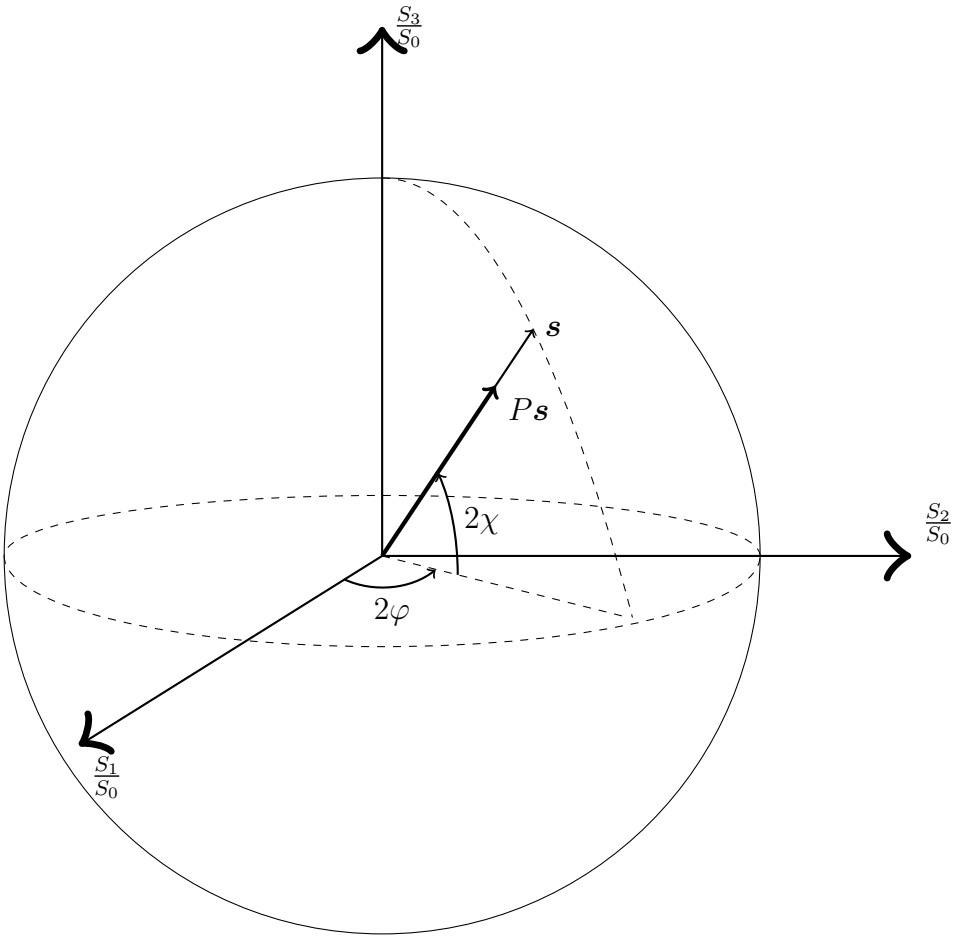


Figure 1.3: Poincaré sphere with completely polarized light $(1, \boldsymbol{s})^T$ and partially polarized light $(1, P\boldsymbol{s})^T$

The end point of $(1, \boldsymbol{s})^T$, which is a completely polarized light, is located on the sphere while a partially polarized light $(1, P\boldsymbol{s})^T$ has an end point within the sphere. The vector of this unpolarized light represents the states with maximum probability during measurement since in reality it fluctuates with time [35]. The evolution of the vector on the Poincaré sphere is described by the Mueller matrix introduced in the next subsection. It helps to

offer a geometrical understanding of the interaction of different optical elements with the input polarized light. Also, the Poincaré sphere describes all possible states of polarization hence its shape can be changed by the Mueller matrices. A visualization of how the shape of Poincaré sphere is influenced by different types of Mueller matrices is available in Chapter 10 in [33].

1.3.2 Mueller matrix

The mapping from input Stokes vector $\boldsymbol{S}_{in}$ to another is described by the Mueller matrix $\mathbf{M}$. $\mathbf{M}$ is a 4×4 matrix which gives the output Stokes vector $\boldsymbol{S}_{out}$ via

$$\boldsymbol{S}_{out} = \mathbf{M}\boldsymbol{S}_{in} \tag{1.40}$$

Its general form is given in [36] as

$$\mathbf{M} = \begin{pmatrix} 1 & \boldsymbol{D}^T \\ \boldsymbol{P} & \mathbf{m} \end{pmatrix} \tag{1.41}$$

where $\boldsymbol{D}^T$ is a 1×3 diattenuation vector and $\boldsymbol{P}$ is a 3×1 polarizance vector. The polarizance vector describes the degree of polarization of the output light when the input light is unpolarized. $\mathbf{m}$ is the 3×3 submatrix. Polar decomposition [37] decomposes this general formalism into its basic polarimetric properties is introduced in Chapter 3.2.1.
Besides this general format, the Mueller matrix can be adapted to represent basic optical elements i.e., diattenuator, retarder, and depolarizer.
For a retarder, its eigen polarization are called the fast axis and slow axis of this retarder. The polarization of the input light will not change if its polarization direction is parallel to either the fast or the slow axis. The refractive index is the lowest along the fast axis direction and largest along the slow axis direction. Hence, the speed of light is the fastest if the polarization of the light is parallel to the fast axis while it is the slowest if the polarization is parallel to the slow axis. This speed difference can generate a phase difference between the phase of components of the input light polarized parallel to the fast axis (δ_{fast}) and parallel to the slow axis (δ_{slow}). Then retardance R of the retarder is defined in Eq.1.42 as the absolute difference between the phase δ_{fast} and the phase δ_{slow}.

$$R = |\delta_{fast} - \delta_{slow}| \tag{1.42}$$

In Mueller-Stokes formalism, the Mueller matrix of a retarder $\mathbf{M}_R$ is a unitary matrix which means the output Stokes vector does not change in total intensity and the degree of

polarization. Its eigenvectors are

$$\mathbf{M}_R \begin{pmatrix} 1 \\ \pm\hat{\boldsymbol{R}} \end{pmatrix} = \begin{pmatrix} 1 \\ \pm\hat{\boldsymbol{R}} \end{pmatrix} \tag{1.43}$$

$(1, \pm\hat{\boldsymbol{R}})$ represents the fast and slow axes of the retarder. The effect of the retardance matrix on the polarized light is equivalent to a rotation matrix on the Poincaré sphere. As shown in Fig.1.4, the retarder rotates the input Stokes vector about the fast axis clockwise of an angle equivalent to the retardance R of the retarder.

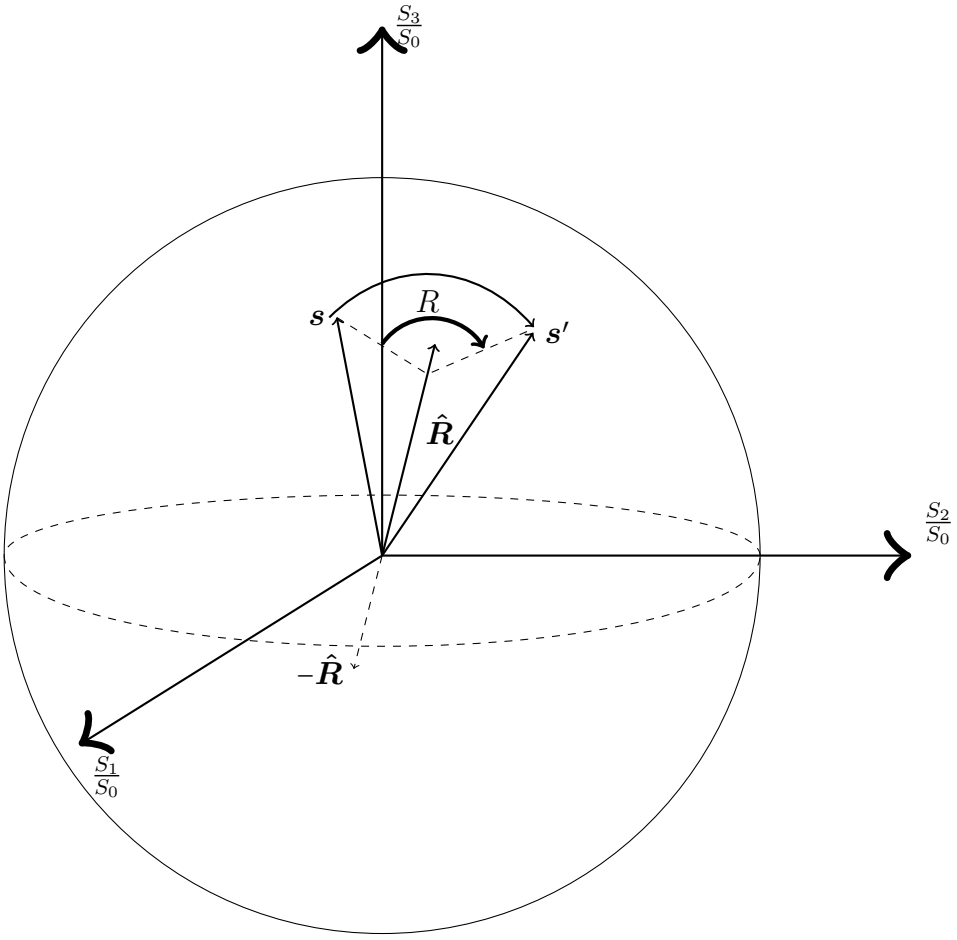


Figure 1.4: A general retarder rotating the Stokes vector $(1, \boldsymbol{s})$ about the fast axis $\hat{\boldsymbol{R}}$ clockwise at an angle R (same as the retardance value of the retarder) to the final Stokes vector $(1, \boldsymbol{s})$. $-\hat{\boldsymbol{R}}$ is the slow axis of the retarder

A general retarder can be seen as a combination of linear retarder and a circular retarder (also called an optical rotator). For a linear retarder, its fast axis lies in the $\frac{S_1}{S_0} - \frac{S_2}{S_0}$ plane. Hence, the rotation about the fast axis of the linear retarder changes both the azimuth and ellipticity of the input polarized light. For a circular retarder, its fast axis is the $\frac{S_3}{S_0}$ axis. Rotating about this axis only changes the azimuth while the ellipticity remains unchanged. Hence, the circular retarder only rotates the light fields without changing the degree of circular polarization in the input light.

A diattenuator can be deemed as a partial polarizer [38] whose transmission is not equal along orthogonal directions. It is represented by the diattenuation matrix $\mathbf{M}_D$ whose eigenvectors $(1\pm, \hat{\boldsymbol{D}})$ are along the directions on the Poincaré sphere with the highest and lowest transmission.

$$\mathbf{M}_D \begin{pmatrix} 1 \\ \pm\hat{\boldsymbol{D}} \end{pmatrix} = T_u(1 \pm D) \begin{pmatrix} 1 \\ \pm\hat{\boldsymbol{D}} \end{pmatrix} \tag{1.44}$$

where T_u is the transmission of unpolarized light and D is the diattenuation of the matrix $\mathbf{M}_D$. $(1, \hat{\boldsymbol{D}})$ is the transmission axis of $\mathbf{M}_D$. As shown in Eq.1.44, the highest transmission T_{max} and lowest transmission T_{min} of $\mathbf{M}_D$ are given as

$$T_{max} = T_u(1 + D) \qquad T_{min} = T_u(1 - D) \tag{1.45}$$

These two values are also the eigenvalues of $\mathbf{M}_D$. From Eq.1.44, the diattenuation value can also be written as

$$D = \frac{T_{max} - T_{min}}{T_{max} + T_{min}} \tag{1.46}$$

Once the T_{min} is close to zero, the diattenuator becomes a polarizer that only transmits light parallel to the transmission axis. If the Stokes vector of the input light is parallel or perpendicular to the diattenuation axis, the state of polarization of the exiting light remains unchanged. In other cases, the output polarized light always lies in the same plane defined by the input polarized light and the transmission axis [37] on the Poincaré sphere. This output vector will lean towards the transmission axis compared to the input.

The Mueller matrix of a depolarizer $\mathbf{M}_\Delta$ imposes a depolarization effect on the input polarized light. To quantify the depolarization ability of $\mathbf{M}_\Delta$, the parameter depolarization in Eq.1.47 is used.

$$\Delta = 1 - \frac{|\mathrm{tr}(\mathbf{M}_\Delta) - 1|}{3} \tag{1.47}$$

The higher Δ, the higher the depolarization ability of the material. Although $\mathbf{M}_\Delta$ represents a depolarizer, unpolarized light can be polarized after transmitting through this depolarizer if it has a nonzero polarizance [37].

1.4 Mueller matrix polarimetry and PSOCT

1.4.1 Basic Mueller matrix polarimeter

To measure the Mueller matrix of the sample, the Mueller matrix polarimeter is used to conduct the Mueller matrix polarimetry (MMP). A Mueller matrix polarimeter is composed of a polarization state generator (PSG) and a polarization state analyzer (PSA). The PSG converts the unpolarized input light to polarized light for interaction with the sample. If the PSG is composed of a horizontal polarizer followed by a quarter wave plate (QWP) whose fast axis is at θ_g, then the Stokes vector generated by this PSG can be written as

$$\boldsymbol{S}_G(\theta_g) = \frac{1}{2}(1, \cos^2(2\theta_g), \cos(2\theta_g)\sin(2\theta_g), \sin(2\theta_g))^T \tag{1.48}$$

This can be computed from consecutive multiplication of the unpolarized light $(1,0,0,0)^T$, the Mueller matrix of a horizontal linear polarizer $\mathbf{M}_{LP0}$ and a QWP $\mathbf{M}_{QWP}$. $\mathbf{M}_{LP0}$ and $\mathbf{M}_{QWP}$ at θ are listed as [39]

$$\mathbf{M}_{LP0} = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.49}$$

$$\mathbf{M}_{QWP} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2(2\theta) & \cos(2\theta)\sin(2\theta) & -\sin(2\theta) \\ 0 & \cos(2\theta)\sin(2\theta) & \sin^2(2\theta) & \sin(2\theta) \\ 0 & \sin(2\theta) & -\cos(2\theta) & 0 \end{pmatrix} \tag{1.50}$$

The interaction with the sample Mueller matrix $\mathbf{M}_s$ yields the output polarized light as

$$\boldsymbol{S}_{out} = \mathbf{M}_s \boldsymbol{S}_G(\theta_g) \tag{1.51}$$

The output polarized light is then measured by the PSA for computing this Stokes vector. Since the camera can only measure the intensity of the output light, one needs 4 independent measurements to determine all the elements in $\boldsymbol{S}_{out}$. For a PSA composed of a QWP whose fast axis is at θ_a followed by a horizontal polarizer, one can represent one measurement by a row vector $(\boldsymbol{S}_A(\theta_a))^T$ as

$$(\boldsymbol{S}_A(\theta_a))^T = \frac{1}{2}(1, \cos^2(2\theta_a), \cos(2\theta_a)\sin(2\theta_a), \sin(2\theta_a) \tag{1.52}$$

So the 4 measured intensities (I_1, I_2, I_3, I_4) are measured by 4 measurements conducted at 4 different angles $(\theta_{a_1}, \theta_{a_2}, \theta_{a_3}, \theta_{a_4})$ which are the fast axis θ_a of the QWP. The process can be written as

$$\begin{pmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{pmatrix} = \begin{pmatrix} (\boldsymbol{S}_A(\theta_{a_1}))^T \\ (\boldsymbol{S}_A(\theta_{a_2}))^T \\ (\boldsymbol{S}_A(\theta_{a_3}))^T \\ (\boldsymbol{S}_A(\theta_{a_4}))^T \end{pmatrix} \boldsymbol{S}_{out} \tag{1.53}$$

The measurement of one output Stokes vector cannot infer all the elements in the Mueller matrix. To compute the Mueller matrix, 4 different input polarized light states have to be generated using four different angles $(\theta_{g_1}, \theta_{g_2}, \theta_{g_3}, \theta_{g_4})$ in the QWP of the PSG and each measured fully by the PSA. The measured intensity $I_{g_m a_n}$ under a pair of PSG: $\boldsymbol{S}_G(\theta_{g_m})$ $(m \in \{1,2,3,4\})$ and PSA: $(\boldsymbol{S}_A(\theta_{a_n}))^T$ $(n \in 1,2,3,4)$ is given as

$$I_{g_m a_n} = (\boldsymbol{S}_A(a_n))^T \mathbf{M}_s \boldsymbol{S}_G(g_m) \tag{1.54}$$

For simplicity, PSG matrix $\mathbf{M}_G$ and PSA matrix $\mathbf{M}_A$ in Eq.1.55,1.56 can be created to express Eq.1.54 in matrix formalism.

$$\mathbf{M}_G = (\boldsymbol{S}_G(\theta_{g_1}), \boldsymbol{S}_G(\theta_{g_2}), \boldsymbol{S}_G(\theta_{g_3}), \boldsymbol{S}_G(\theta_{g_4})) \tag{1.55}$$

$$\mathbf{M}_A = \begin{pmatrix} (\boldsymbol{S}_A(\theta_{a_1}))^T \\ (\boldsymbol{S}_A(\theta_{a_2}))^T \\ (\boldsymbol{S}_A(\theta_{a_3}))^T \\ (\boldsymbol{S}_A(\theta_{a_4}))^T \end{pmatrix} \tag{1.56}$$

Then, the intensity $I_{g_m g_a}$ in Eq.1.54 can form an intensity matrix which is equivalent to

$$\begin{pmatrix} I_{g_1 a_1} & I_{g_2 a_1} & I_{g_3 a_1} & I_{g_4 a_1} \\ I_{g_1 a_2} & I_{g_2 a_2} & I_{g_3 a_2} & I_{g_4 a_2} \\ I_{g_1 a_3} & I_{g_2 a_3} & I_{g_3 a_3} & I_{g_4 a_3} \\ I_{g_1 a_4} & I_{g_2 a_4} & I_{g_3 a_4} & I_{g_4 a_4} \end{pmatrix} = \mathbf{M}_A \mathbf{M}_s \mathbf{M}_G \tag{1.57}$$

Representing the intensity matrix as $\mathbf{I}$, the Mueller matrix can be computed as

$$\mathbf{M}_s = (\mathbf{M}_A)^{-1} \mathbf{I} (\mathbf{M}_G)^{-1} \tag{1.58}$$

It is clear that $\mathbf{M}_A$ and $\mathbf{M}_G$ should be invertible for computing M_s in Eq.1.58. Hence, columns in M_G and rows in M_A should be linearly independent. Consider the 4 Stokes vectors constructing the columns of M_G when the QWP in the PSG is replaced by a variable phase retarder. Since the fast axis of the phase retarder in the PSG is fixed, the

4 Stokes vectors will share the same plane on the Poincaré sphere, resulting in 4 linearly dependent Stokes vectors. The resulting M_G is not invertible. Hence, even if the PSG generates four different Stokes vector, the measurement of the Mueller matrix can still fail. So the variable phase retarder has to change its fast axis after generating 2 different Stokes vectors using the same fast axis. Thus, Mueller matrix polarimeters using phase modulators require at least 2 phase modulators in the PSG and 2 phase modulator in the PSA. The final 4 Stokes vectors can then form a tetrahedra on the Poincaré sphere. As shown in [40, 41], the larger the volume of this tetrahedra, the better the stability of the measurement result when influenced by system errors.
Rotating the QWP in the PSA to 4 angles can be time consuming. One setup called division of amplitude polarimeter [42, 43] divides the light into 4 channels so that one can implement 4 PSAs and 4 cameras to measure the Stokes vector fully at the same instance. Thus, it can help to reduce the Mueller matrix polarimetry to 4 rotations of the QWP in the PSG instead of 16 rotations in the PSG and PSA. However, once dividing the light intensity 4 times, the signal to noise ratio (SNR) of the image will go down. Additionally, the cost of the device can also be more expensive because of increased number of QWPs and cameras.

1.4.2 Dual rotating QWP polarimeter

Most Mueller matrix polarimeters are either polarimeters with rotating elements or polarimeters with phase modulators [38].
In a rotating elements setup, the rotating elements can be either QWPs or polarizers. The first setup of the dual rotating QWP polarimeter is given in [44]. A simple illustration of the dual rotating QWP setup is shown here. First, the setup is given in Fig.1.5.

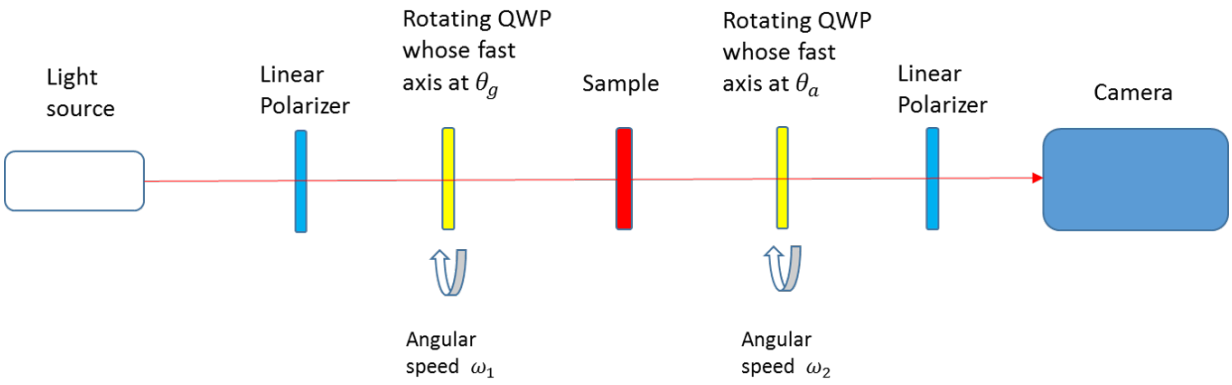


Figure 1.5: Experimental setup for dual rotating QWP polarimeter.

As shown in Eq.1.53,1.55, the measured intensity I when the angle of the QWP in PSG is θ_g and the angle of the QWP in PSA is θ_a

$$I = \frac{1}{4}(1, \cos^2(2\theta_a), \cos(2\theta_a)\sin(2\theta_a), -\sin(2\theta_a)) \begin{pmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \cos^2(2\theta_g) \\ \cos(2\theta_g)\sin(2\theta_g) \\ \sin(2\theta_g) \end{pmatrix} \tag{1.59}$$

In the frequency modulation given in [44], the angular speed of rotation of the QWP is ω_1 in PSG and ω_2 in PSA. If the orientation of the fast axes of the QWPs in the PSG and PSA before rotation is at zero, at time t, the angle of the fast axis of the QWP in the PSA is $\omega_1 t$ and the fast axis of the QWP in PSA is $\omega_2 t$. Hence, the intensity in Eq.1.59 can be

expressed in terms of time as

$$
\begin{aligned}
I(t) =& \frac{1}{4}(m_{00} + \frac{1}{2}m_{01} + \frac{1}{2}m_{10} + \frac{1}{4}m_{11} \\
&+ (m_{03} + \frac{1}{2}m_{13})\sin(2\omega_1 t) \\
&- (m_{30} + \frac{1}{2}m_{31})\sin(2\omega_2 t) \\
&+ (\frac{1}{2}m_{02} + \frac{1}{4}m_{12})\sin(4\omega_1 t) + (\frac{1}{2}m_{01} + \frac{1}{4}m_{11})\cos(4\omega_1 t) \\
&+ (\frac{1}{2}m_{20} + \frac{1}{4}m_{21})\sin(4\omega_2 t) + (\frac{1}{2}m_{10} + \frac{1}{4}m_{11})\cos(4\omega_2 t) \\
&- \frac{1}{2}m_{33}\cos(2\omega_1 t - 2\omega_2 t) \\
&+ \frac{1}{2}m_{33}\cos(2\omega_1 t + 2\omega_2 t) \\
&+ \frac{1}{4}m_{13}\sin(2\omega_1 + 4\omega_2 t) - \frac{1}{4}m_{23}\cos(2\omega_1 t + 4\omega_2 t) \\
&- \frac{1}{4}m_{31}\sin(4\omega_1 t + 2\omega_2 t) + \frac{1}{4}m_{32}\cos(4\omega_1 t + 2\omega_2 t) \\
&+ \frac{1}{4}m_{13}\sin(2\omega_1 t - 4\omega_2 t) + \frac{1}{4}m_{23}\cos(2\omega_1 t - 4\omega_2 t) \\
&+ \frac{1}{4}m_{31}\sin(4\omega_1 t - 2\omega_2 t) - \frac{1}{4}m_{32}\cos(4\omega_1 t - 2\omega_2 t) \\
&+ \frac{1}{8}(m_{12} - m_{21})\sin(4\omega_1 t - 4\omega_2 t) + \frac{1}{8}(m_{11} + m_{22})\cos(4\omega_1 t - 4\omega_2 t) \\
&+ \frac{1}{8}(m_{12} + m_{21})\sin(4\omega_1 t + 4\omega_2 t) + \frac{1}{8}(m_{11} - m_{22})\cos(4\omega_1 t + 4\omega_2 t))
\end{aligned}
\tag{1.60}
$$

Hence, the intensity signal is a function of 13 different frequencies (including 0). These coefficients can be computed using discrete Fourier transform of the signal $I(t)$. Representing each coefficient before sine function as $a(f(\omega))$ and coefficient before cosine function as $b(f(\omega))$, where the $f(\omega)$ is the frequency combination of ω_1 and ω_2. Set $\frac{1}{4}(m_{00} + \frac{1}{2}m_{01} + \frac{1}{2}m_{10} + \frac{1}{4}m_{11})$ as a_0. Then each element in the Mueller matrix can be

written as

$$
\begin{aligned}
m_{00} =& a_0 - \frac{1}{4}[4b(4\omega_1 t) - 4(b(4\omega_1 t - 4\omega_2 t) + b(4\omega_1 t + 4\omega_2 t))] \\
& - \frac{1}{4}[4b(4\omega_2 t) - 4(b(4\omega_1 t - 4\omega_2 t) + b(4\omega_1 t + 4\omega_2 t))] - b(4\omega_1 t - 4\omega_2 t) - b(4\omega_1 t + 4\omega_2 t) \\
m_{01} =& \frac{1}{2}[4b(4\omega_1 t) - 4(b(4\omega_1 t - 4\omega_2 t) + b(4\omega_1 t + 4\omega_2 t))] \\
m_{02} =& \frac{1}{2}[4a(4\omega_1 t) - 4(a(4\omega_1 t + 4\omega_2 t) + a(4\omega_1 t - 4\omega_2 t))] \\
m_{03} =& \frac{1}{2}[2a(2\omega_1 t) - 4a(2\omega_1 t + 4\omega_2 t)] \\
m_{10} =& \frac{1}{2}[4b(4\omega_2 t) - 4(b(4\omega_1 t - 4\omega_2 t) + b(4\omega_1 t + 4\omega_2 t))] \\
m_{11} =& 4(b(4\omega_1 t - 4\omega_2 t) + b(4\omega_1 t + 4\omega_2 t)) \\
m_{12} =& 4(a(4\omega_1 t + 4\omega_2 t) + a(4\omega_1 t - 4\omega_2 t)) \\
m_{13} =& 4a(2\omega_1 t + 4\omega_2 t) \\
m_{20} =& \frac{1}{2}[4a(4\omega_2 t) - 4(a(4\omega_1 t + 4\omega_2 t) - a(4\omega_1 t - 4\omega_2 t))] \\
m_{21} =& 4(a(4\omega_1 t + 4\omega_2 t) - a(4\omega_1 t - 4\omega_2 t)) \\
m_{22} =& 4(b(4\omega_1 t - 4\omega_2 t) - b(4\omega_1 t + 4\omega_2 t)) \\
m_{23} =& -4b(2\omega_1 t + 4\omega_2 t) \\
m_{30} =& -\frac{1}{2}[-2a(\omega_2 t) + 4a(4\omega_1 t + 2\omega_2 t)] \\
m_{31} =& -4a(4\omega_1 t + 2\omega_2 t) \\
m_{32} =& 4b(4\omega_1 t + 2\omega_2 t) \\
m_{33} =& 2b(2\omega_1 t + 2\omega_2 t)
\end{aligned}
\tag{1.61}
$$

1.4.3 PSOCT

Another type of instrument capable of quantifying the polarization properties of the sample is PSOCT. It is capable of measuring depth resolved polarimetric properties of the sample. Application of PSOCT includes measuring birefringence of *ex vivo* tissue samples for 3D structural analysis [45, 46, 47, 48], *in vivo* retinal imaging for thickness measurement and layer identification [49, 50], and *in vivo* skin imaging for scar assessment [51].

PSOCT is built based on optical coherence tomography (OCT) systems. The OCT system uses the interference of light reflected from different depths with the light reflected from a reference mirror to probe the depth resolved reflectivity of the sample. The depth resolved reflectivity is then used to create a 3D model of the sample. There are two types of OCT imaging modalities. One is the Time Domain OCT (TDOCT) which achieves the depth information in the sample arm by moving the reference arm. Another is the Fourier Domain OCT (FDOCT) which is based on the Fourier analysis of the output spectrum data. PSOCT systems are mostly based on the latter modality which is significantly faster than the former. FDOCTs are either Spectral Domain OCT (SDOCT) which uses a broadband light source or Swept Source OCT (SSOCT) which scans across a certain spectrum.

There are a few different configurations in PSOCT [52]. One is the setup called the Hee-Hiztenberger type of PSOCT shown in Fig.1.6.

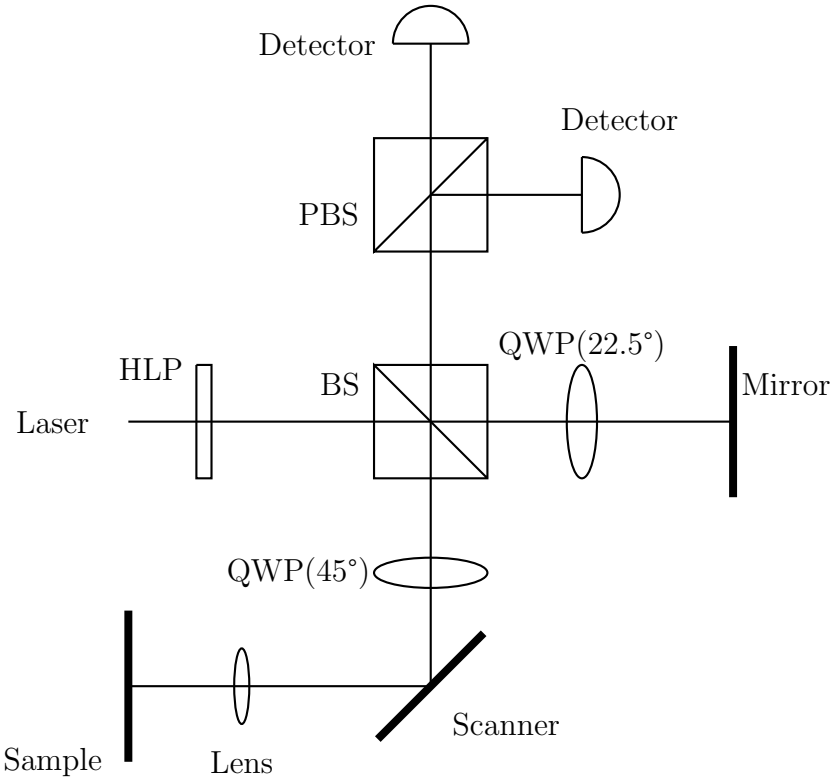


Figure 1.6: Simple diagram of Hee-Hiztenberger PSOCT. BS: beam splitter, PBS: polarization beam splitter, HLP: horizontal linear polarizer. Components in reference arm: QWP (fast axis at 22.5°) and mirror. Components in sample arm: QWP (fast axis at 45°), scanner, lens, and sample.

This type of system deems the sample as a linear retarder. By measuring the phase change of the output polarized light, one can estimate the depth resolved linear retardance. The first system to implement this assumption was presented by Hee, *et al.* [53] and further developed in [54, 55]. The input light is first polarized by the polarizer and split into the reference arm and the sample arm. The sample arm is the light path in Fig.1.6 with a sample located at the end. The reference arm is the light path with a mirror located in the end. In each arm, there is a QWP oriented 22.5° to turn the input horizontal linearly polarized light to elliptical polarized light. The reflected light from the reference arm goes through the QWP again which returns a polarized light along the −45°. In the sample arm, the fast axis of the QWP is at 45° to produce circularly polarized light to interact with sample. If the sample is a mirror, the reflected light after a second path of the QWP is a

vertically polarized light. Hence, the returned light from the sample arm will only interfere with the vertical component of the output light from the sample arm. However, if the sample is not a mirror but a linear retarder, the returned light from the sample arm will have components in both horizontal and vertical channels. The interference of the light from the sample and reference arm is then split into horizontal and vertical channels by a polarization beam splitter for measuring the spectrum. The depth resolved intensity in the horizontal channel $I_H(z)$ and vertical channel $I_V(z)$ are computed by applying inverse Fourier Transform of the spectrum and their expression can be written as [52]

$$I_H(z) \propto R(z)\sin^2(\delta(z)) \tag{1.62}$$
$$I_V(z) \propto R(z)\cos^2(\delta(z)) \tag{1.63}$$

where $R(z)$ is the reflectivity at z and $\delta(z)$ is the phase retardation at z. Hence, the depth resolved linear birefringence $\Delta n(z)$ is computed as the gradient of the depth resolved linear retardance as

$$\begin{aligned}\Delta n(z) =& \frac{1}{k_0}\frac{d\delta(z)}{dz} \\ =& \frac{1}{k_0}\frac{d}{dz}\Big(\arctan\sqrt{\frac{I_H}{I_V}}\Big)\end{aligned} \tag{1.64}$$

Hence, compared to a traditional OCT system which only acquires the depth resolved reflectivity, the Hee-Hiztenberger type PSOCT also acquires the birefringence and its orientation at each depth. However, the price of the device can also increase since it requires one more spectrometer and camera for simultaneous determination of the horizontal and vertical component of the output interference signal.

Another method called the Jones matrix OCT has been proposed in [56]. It measures the depth resolved Jones matrix of the sample. So it provides 3 parameters in diattenuation and the circular retardance in addition to the two parameters acquiring using the Hee-Hiztenberger's method. Both the Hee-Hiztenberger's setup and Jones matrix OCT are based on simultaneous measurements to infer the polarimetric properties. However, PSOCT which uses Mueller-Stokes formalism [57, 58] requires multiple numbers of measurements at different times to fully determine the depth resolve Mueller matrix, resulting in significant difficulties in *in vivo* measurements. Hence, although providing a more thorough estimate of the polarimetric properties of the sample, most PSOCT systems are not based on Mueller-Stokes formalism.

1.5 Decomposition of the Mueller matrix for polarimetric parameters

The Mueller matrix acquired from Mueller matrix polarimetry requires decomposition to extract the polarimetric parameters. There are many different methods for decomposing the Mueller matrix, most of which belong to three major categories of decomposition methods including serial, parallel, and differential decomposition [33]. The serial decomposition considers the medium as a sequence of optical elements listed in the light path so that the input light transmits these elements one after another. On the other hand, parallel decomposition deems the medium as many independent optical elements listed in parallel so that the input light is split into these independent pathways to and recombined at the output. Another decomposition method is the differential decomposition. The differential decomposition decomposes the medium as a continuous homogenous medium with each infinitesimal slab having the same polarization properties. Since the serial and differential decompositions are more widely used than parallel, a general introduction of the serial and differential decompositions are given in this section. However, only the polar decomposition, which is a type of the serial decomposition, is used in this thesis.

1.5.1 Serial decomposition

The most notable member of the serial decomposition is the polar decomposition proposed by Lu and Chipman in [37]. The decomposition deems the medium as an consecutive multiplication of three matrices introduced in section 1.3.2, the depolarization matrix $\mathbf{M}_\Delta$, the retardance matrix $\mathbf{M}_R$, and diattenuation matrix $\mathbf{M}_D$. The Mueller matrix is written as

$$\mathbf{M} = \mathbf{M}_\Delta \mathbf{M}_R \mathbf{M}_D \tag{1.65}$$

The polar decomposition illustrated in the original paper [37] starts from decomposing $\mathbf{M}_D$ first, followed by $\mathbf{M}_\Delta$ and $\mathbf{M}_R$. The detailed description of the polar decomposition is illustrated in Chapter 3.2.1.

Another serial decomposition method is the symmetric decomposition [59, 60]. One problem of the polar decomposition is that it lacks symmetry. These three matrices can be placed in different orders. For example, instead of $\mathbf{M}_\Delta \mathbf{M}_R \mathbf{M}_D$, decomposing it in the order of $\mathbf{M}_R \mathbf{M}_\Delta \mathbf{M}_D$ is also viable. This will also generate slightly different results. Hence, symmetrical decomposition provides another way of decomposing the Mueller matrix to acquire a definite solution. First proposed in [36], a Mueller matrix can be written as

$$\mathbf{M} = \mathbf{M}_{J2} \mathbf{M}_\Delta \mathbf{M}_{J1} \tag{1.66}$$

$\mathbf{M}_{J2}$ and $\mathbf{M}_{J1}$ are two pure Mueller matrices with no depolarization. $\mathbf{M}_{\Delta d}$ is a diagonal depolarizer. In [59], this form of decomposition is established as decomposing the Mueller matrix into a symmetric form as

$$\mathbf{M} = \mathbf{M}_{D2}\mathbf{M}_{R2}\mathbf{M}_{\Delta}\mathbf{M}_{R1}\mathbf{M}_{D1} \tag{1.67}$$

where on both sides of the depolarizer are one retardance matrix and one diattenuation matrix. The decomposition first decomposes the outside diattenuation matrices, then the retardance matrices. Hence, a Mueller matrix decomposed by symmetric decomposition will have two diattenuation values and two retardance values. However, comparing to the polar decomposition, the symmetric decomposition requires more computation steps to fully decompose a Mueller matrix. Additionally, the sample diattenuation and retardance are represented by a combination of two values which can make the interpretation of these values difficult.

1.5.2 Differential decomposition

One drawback of the serial decomposition is its arbitrary formalism. The natural sample is not a simple combination of several optical elements followed by one another. The assumption in differential decomposition is to treat the sample as a homogeneous medium where every thickness along the light path is the same in optical properties. Suppose the Mueller matrix from 0 to z is $\mathbf{M}(0, z)$ and from 0 to $z + \Delta z$ is $\mathbf{M}(0, z + \Delta z)$. The medium in between point z and $z + \Delta z$ is $\mathbf{M}(z, z + \Delta z)$, then

$$\mathbf{M}(0, z + \Delta z) = \mathbf{M}(z, z + \Delta z)\mathbf{M}(0, z) \tag{1.68}$$

Subtracting both sides with $\mathbf{M}(0, z)$ and dividing by Δz gives

$$\frac{\mathbf{M}(0, z + \Delta z) - \mathbf{M}(0, z)}{\Delta z} = \frac{\mathbf{M}(z, z + \Delta z) - \mathbf{I}}{\Delta z}\mathbf{M}(0, z) \tag{1.69}$$

At the limit of $\Delta z \to 0$, the above equation is rearranged as

$$\mathbf{M}(0, z)^{-1} \lim_{\Delta z \to 0} \frac{\mathbf{M}(0, z + \Delta z) - \mathbf{M}(0, z)}{\Delta z} = \lim_{\Delta z \to 0} \frac{\mathbf{M}(z, z + \Delta z) - \mathbf{I}}{\Delta z} \tag{1.70}$$

The differential Mueller matrix $\mathbf{m}$ is then defined as the right hand side of Eq.1.70. This equation is equivalent to

$$\mathbf{M}(0, z)^{-1} d\mathbf{M}(z) = \mathbf{m} dz \tag{1.71}$$

Integrating from 0 to z gives

$$\ln \mathbf{M}(0,z) = \int_0^z \mathbf{m} dz \tag{1.72}$$

Hence, the differential Mueller matrix can be written as

$$\mathbf{m}z = \ln \mathbf{M}(0,z) \tag{1.73}$$

The specific form of the differential form of the Mueller matrix is given in [61]. The decomposed polarimetric properties can be directly retrieved from the entries of the differential Mueller matrix by separating it into depolarizing matrix $\mathbf{m}_u$ and non depolarizing matrix $\mathbf{m}_m$ as

$$\mathbf{m}_m = \frac{1}{2}(\mathbf{m} - \mathbf{G}\mathbf{m}^T\mathbf{G}) \qquad \mathbf{m}_u = \frac{1}{2}(\mathbf{m} + \mathbf{G}\mathbf{m}^T\mathbf{G}) \tag{1.74}$$

where $\mathbf{G}$ is the diagonal matrix with $(1,-1,-1,-1)$ as its diagonal entries. The depolarization power is solved from the trace of $\mathbf{m}_u$ while the non depolarization part of the Mueller matrix $\mathbf{m}_m$ has the form of

$$\mathbf{m}_m z = z \begin{pmatrix} 0 & LD & LD' & CD \\ LD & 0 & CB & -LB' \\ LD' & -CB & 0 & LB \\ CD & LB' & -LB & 0 \end{pmatrix} \tag{1.75}$$

The elements LB, LB' are the linear birefringence along horizontal-vertical and ±45° axes while CB is the circular birefringence. Similarly, the elements LD, LD' are the linear dichroism along horizontal-vertical, ±45° axes and CD represents the circular dichroism. The multiplication of the path length z with each term gives the retardance and diattenuation of the medium. Hence, the differential decomposition extracts the retardance and diattenuation from the Mueller matrix by assuming homogeneous property of the medium. However, the natural logarithm of a matrix takes a much longer time to process than polar decomposition which will result in long processing time for computing the properties for large image.

1.5.3 List of parameters used in this thesis

In this thesis, parameters related to diattenuation, polarizance, depolarization and retardance are used. In addition, other parameters including Depolarizaiton index DI, Q metric Q, linear anisotropy and circular anisotropy are used in Chapter 3 and 4.
Depolarization index derived in [62] is similar to the depolarization power Δ introduced in

section 1.3.2, except the depolarization index is not computed from the polar decomposition. The Q metric [63] is also dependent on depolarization but it combines information of diattenuation, polarizance with depolarization. The parameter linear anisotropy A_L and circular anisotropy A_C are given in [64]. These quantities remain unchanged under linear and circular transform. Their computations are listed in here as

$$A_L = \frac{1}{\sqrt{\Sigma}}\sqrt{(m_{01}+m_{10})^2+(m_{02}+m_{20})^2+(m_{23}-m_{32})^2+(m_{13}-m_{31})^2} \tag{1.76}$$

$$A_C = \frac{1}{\sqrt{\Sigma}}\sqrt{(m_{03}+m_{30})^2+(m_{12}-m_{21})^2} \tag{1.77}$$

where Σ is

$$\Sigma = (3m_{00}^2 - m_{11}^2 - m_{22}^2 - m_{33}^2) + 2(m_{01}m_{10} + m_{02}m_{20} + m_{03}m_{30} - m_{23}m_{32} - m_{13}m_{31} - m_{12}m_{21}) \tag{1.78}$$

All the parameters used are summarized in Table 1.1 with their names, symbols, and method of computation listed.

Parameter name	Symbol	Equation	Require **PD**
Horizontal-vertical diattenuation	D_H	$\mathbf{M}_{01}$	No
45°-135° diattenuation	D_{45}	$\mathbf{M}_{02}$	No
Linear diattenuation	D_L	$\sqrt{D_H^2 + D_{45}^2}$	No
Circular diattenuation	D_C	$\mathbf{M}_{03}$	No
Orientation of D_L	D_θ	$\frac{1}{2}\arctan(\frac{D_{45}}{D_H})$	No
Total diattenuation	D	$\sqrt{D_L^2 + D_C^2}$	No
Horizontal-vertical Polarizance	P_H	$\mathbf{M}_{10}$	No
45°-135° Polarizance	P_{45}	$\mathbf{M}_{20}$	No
Linear polarizance	P_L	$\sqrt{P_H^2 + P_{45}^2}$	No
Circular polarizance	P_C	$\mathbf{M}_{30}$	No
Orientation of P_L	P_θ	$\frac{1}{2}\arctan(\frac{P_{45}}{P_H})$	No
Total polarizance	P	$\sqrt{P_L^2 + P_C^2}$	No
Depolarization index	DI	$\sqrt{\frac{1}{3m_{00}^2}(\sum_{i,j=0}^{3} m_{i,j}^2 - m_{00}^2)}$	No
Q metric	Q	$\frac{3DI^2 - D^2}{1+D^2}$	No
Linear anisotropy	A_L	Eq.1.76	No
Circular anisotropy	A_C	Eq.1.77	No
Polarizance of M_D	P_Δ	$\frac{\boldsymbol{P}-\mathbf{m}\boldsymbol{D}}{1-D^2}$	No
Depolarization power	Δ	$1 - \frac{\lvert\mathrm{tr}(M_\Delta)-1\rvert}{3}$	Yes
Total Retardance	R	$\cos^{-1}[\frac{\mathrm{tr}(\mathbf{M}_R)}{2} - 1]$	Yes
Horizontal-vertical retardance	R_H	$\frac{1}{2\sin R}(\mathbf{M}_R(23) - \mathbf{M}_R(32)$	Yes
45°-135° retardance	R_{45}	$\frac{1}{2\sin R}(\mathbf{M}_R(31) - \mathbf{M}_R(13)$	Yes
Linear retardance	R_L	$\sqrt{R_H^2 + R_{45}^2}$	Yes
Circular retardance	R_C	$\frac{1}{2\sin R}(\mathbf{M}_R(12) - \mathbf{M}_R(21)$	Yes
Orientation of R_L	R_θ	$\frac{1}{2}\arctan(\frac{R_{45}}{R_H})$	Yes

Table 1.1: Table of polarimetric parameters used in this thesis with names, symbols, equations given. The fourth column "Require **PD**" indicates whether this parameter requires polar decomposition to solve.

Chapter 2

The birefringence of amyloid deposits in the retina from subjects with Alzheimer's disease

2.1 Introduction

Retinal amyloid deposits, which are presumed to contain Aβ fibrils, possess polarization properties (including retardance) which are significantly different from the surrounding retina [26, 27, 65]. Especially, the linear retardance value is directly linked to the birefringent property of the deposits. As introduced in Section 1.2, the birefringence of senile plaques, which are also presumed to contain Aβ fibrils, have been studied. The birefringence value of unstained senile plaques was reported in [29] using PSOCT. Thus, it would be useful to also measure the birefringence value of the retinal amyloid deposits. To compute the birefringence of the deposits, two parameters need to be measured. One is the linear retardance of the deposit and the other is the thickness of the deposit. Linear retardance can be measured using MMP introduced in Section 1.4. Thickness can be measured using the Confocal Laser Scanning Microscope (CLSM). By combining the two measurements, one can compute the linear birefringence of the presumed Aβ deposits.
In this chapter, the relationship between linear retardance, thickness and the birefringence values of the retinal amyloid deposits are reported. The computed birefringence of the retinal amyloid deposits are then compared with the birefringence values of pure Aβ deposits grown on the glass. The birefringence values of retinal amyloid deposits are also compared with the birefringence of senile plaques [29] and retinal nerve fiber layer (RNFL)

[66, 50, 67]. The results presented here for the retardance, thickness and birefringence values of retinal amyloid deposits demonstrate the feasibility of *in vivo* imaging of the deposits without a contrast agent.

2.2 Method

2.2.1 Sample preparation

2.2.1.1 Retinal samples

Eyes were donated and collected post-mortem in compliance with the Declaration of Helsinki from 4 individuals whose brain pathology including amyloid deposits was consistent with Alzheimers disease pathology at an intermediate or high level under NIA-AA guidelines [68]. Eyes were fixed in 10 % formalin for 24 hours and the retinas are flat mounted. Samples were then stained with Thioflavin-S for amyloid and DAPI for nuclear material. In total, 31 deposits from these 4 individuals were studied.

2.2.1.2 Pure amyloid samples

Pure amyloid deposits were prepared for comparison with retinal amyloid deposits. Aβ(1-42), Ultra pure, NaOH were purchased from rPeptide were used [69]. To mimic the environment of human body, 1ml of 10mM HEPES and 150mM NaCl (pH of 7.4) was added to 0.5mg of Aβ(1-42) and incubated at 37 °C for 89 hours without shaking. After the incubation, 500 μl of filtered 1.25 mM Thioflavin-T in 50 nM PBS were added to the Aβ(1-42) solution. Taking 100 μl alioquots onto 10 slides and letting it sit for 5 minutes. The slides were then rinsed with 50 μl of culture grade water (x3) and dried with a gentle stream of compressed N_2. To match the number of deposits in retina, a total of 31 pure deposits were studied.

2.2.2 Instrument and experiments

2.2.2.1 Confocal fluorescence microscopy

The 3D structure of the samples was examined in a Carl Zeiss LSM 700 confocal laser scanning microscope by using 405 nm and 488 nm wavelengths to excite DAPI and Tioflavin-S

respectively for the retinal samples. For pure deposits, a 488 nm light source was used to excite the Thioflavin-T. A sequence of *en face* images were acquired as a function of depth. These images were used to create a 3D model for computing thickness (see Section 2.2.4).

2.2.2.2 Mueller matrix polarimetry

A Nikon upright transmission fluorescent microscope was modified by adding a PSG before the light interacts with the sample and a PSA after the light was transmitted through the sample. The system setup is illustrated in Fig.2.1.

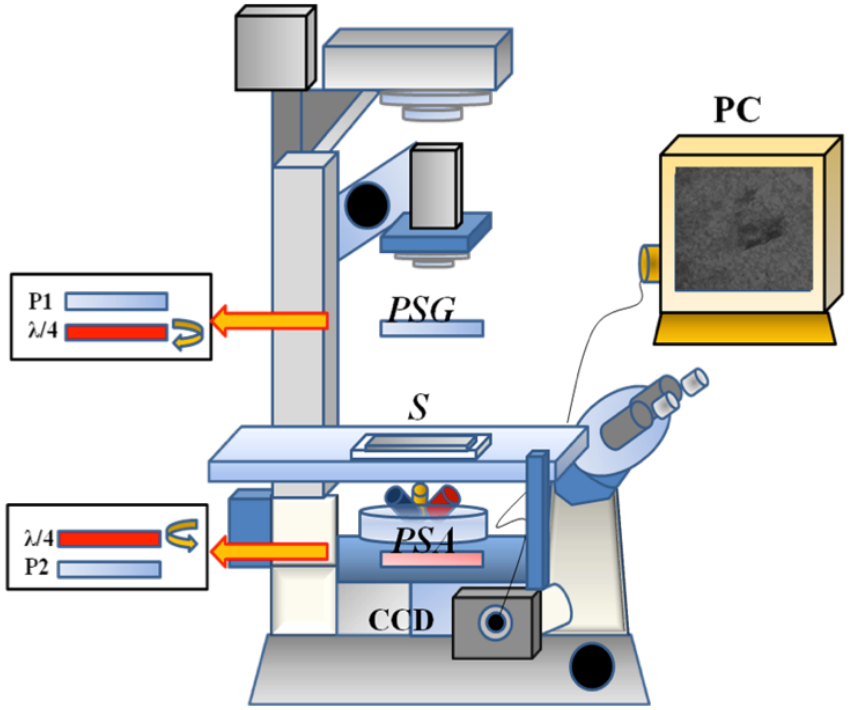


Figure 2.1: Setup of polarization microscope for Mueller matrix. The PSG is composed of a linear polarizer (P1) followed by a QWP ($\lambda/4$). The PSA is composed of a QWP ($\lambda/4$) followed by another linear polarizer (P2)

This is one of the basic configurations introduced in Chapter 1.4. In this microscope, the QWP in the PSG was rotated to 4 angles (45°,0°,−60°,−30°) and four different polarized states of light were generated. The PSA was rotated to the same 4 angles and analyzed the output light after it interacted with the sample by recording a total of 16 images. After

image acquisition, registration was conducted to align the images (see section 2.2.3.1). The recorded 16 images can be written as Eq.2.1 by replacing the two angles in Eq.1.57.

$$\begin{pmatrix} I_{45,45} & I_{00,45} & I_{60,45} & I_{30,45} \\ I_{45,00} & I_{00,00} & I_{60,00} & I_{30,00} \\ I_{45,60} & I_{00,60} & I_{60,60} & I_{30,60} \\ I_{45,30} & I_{00,30} & I_{60,30} & I_{60,60} \end{pmatrix} \tag{2.1}$$

Then, using Eq.1.58, the Mueller matrix of the sample can be computed.

2.2.2.3 Polar decomposition

The Mueller matrix computed from the measurement described carries information about the interaction of the medium with polarized light of any states (its polarimetric properties). These properties can be extracted from the Mueller matrix. Introduced in section 1.5.1, the polar decomposition [37] decomposes the Mueller matrix into a serial multiplication of depolarization matrix $\mathbf{M}_{\Delta}$, retardance matrix $\mathbf{M}_{R}$, and diattenuation $\mathbf{M}_{D}$. 3 parameters were computed and investigated in this chapter: linear retardance LR, total diattenuation D, and depolarization power DP. The effectiveness of these polarization properties in differentiating the deposit and its surrounding background (retina or glass) is give by the Michelson contrast

$$\text{Contrast} = \frac{\mu_{\text{deposit}} - \mu_{\text{background}}}{\mu_{\text{deposit}} + \mu_{\text{background}}} \tag{2.2}$$

where μ represents the mean of the polarimetric parameter in the area.

2.2.3 Image processing

2.2.3.1 Image registration

The 16 images taken using Mueller matrix polarimetry were registered before computing the Mueller matrix from the images. Phase correlation based registration [70] was built inside the image acquisition system to register the images directly after taking them. The intensity based registration [71] was used to align the computed LR from MMP and thickness maps from CLSM for computing the linear birefringence of the deposit on each pixel. Thus, the linear birefringence can be computed by

$$\Delta n = \frac{\lambda}{2\pi} \times \frac{\text{LR}}{\text{thickness}} \tag{2.3}$$

2.2.4 Image segmentation

Image segmentation was performed to separate the deposit from its background in 2D polarization and 3D and CLSM images so that parameters such as the contrast defined in Eq.2.2 could be calculated. The Markov random field (MRF) [72] method was used on the computed Mueller matrix images to separate retinal deposits from background retina and pure deposits from background glass. Methods based on MRF have been used in either segmenting the Mueller matrix images [73, 74] or recovering images taken with a Stokes polarimeter [75, 72]. The advantage of the MRF is that it considers the neighboring pixels when assigning a pixel to one or the other class and it can be fully automated.
To images taken with 488nm excitation giving thioflavin fluorescence, an active contour based Chen-Vese level set method [71, 76] for segmenting the 3D voxels was applied. This method can detect holes inside a 3D volume. After the segmentation, the segmented voxels were summed along the depth direction to construct the *en face* thickness. The thickness map was then registered with the segmented linear retardance images of the deposits.

2.3 Results

2.3.1 Segmentation

The segmentation of the Mueller matrix images of the retinal and pure deposits using MRF are displayed in Fig.2.2.

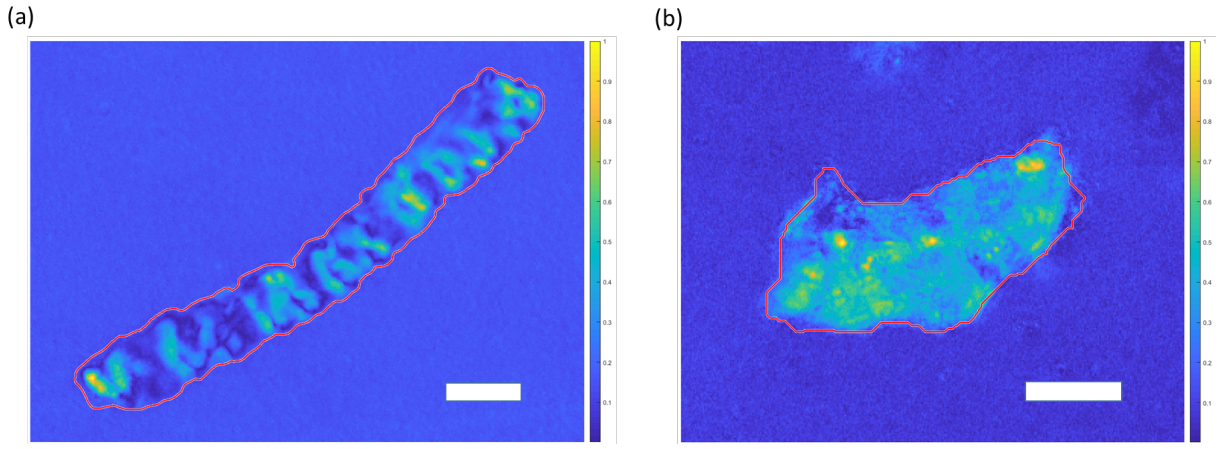


Figure 2.2: Segmentation result of the Mueller matrix images displayed on the normalized linear retardance image of (a) a retinal deposit and (b) a pure deposit. Red contours in both images represent the boundary of the segmentation mask. Scale bar: 20 μm.

As shown in Fig.2.2, the contours closely follow the boundary of the retinal and pure deposits. MRF successfully segmented all 31 retinal and 31 pure deposits.

2.3.2 Thickness and registration

The multiple stacks of images of the retinal and pure deposits generated by the CLSM are displayed using ZEN 2.3 lite [77] in Fig.2.3,(a),(c). In retinal samples, Thioflavin S binds the β sheet structure of the amyloid deposits [78] and the DAPI stains the nucleic acid of the ganglion cells. The thickness maps of the deposits reconstructed form the Thioflavin signal, as shown in Fig.2.3.(b), (d). The pure deposit in Fig.2.3.(d) has an average thickness of 4.8 μm while the retinal deposit in Fig.2.3.(b) has an average thickness of 2.5 μm.

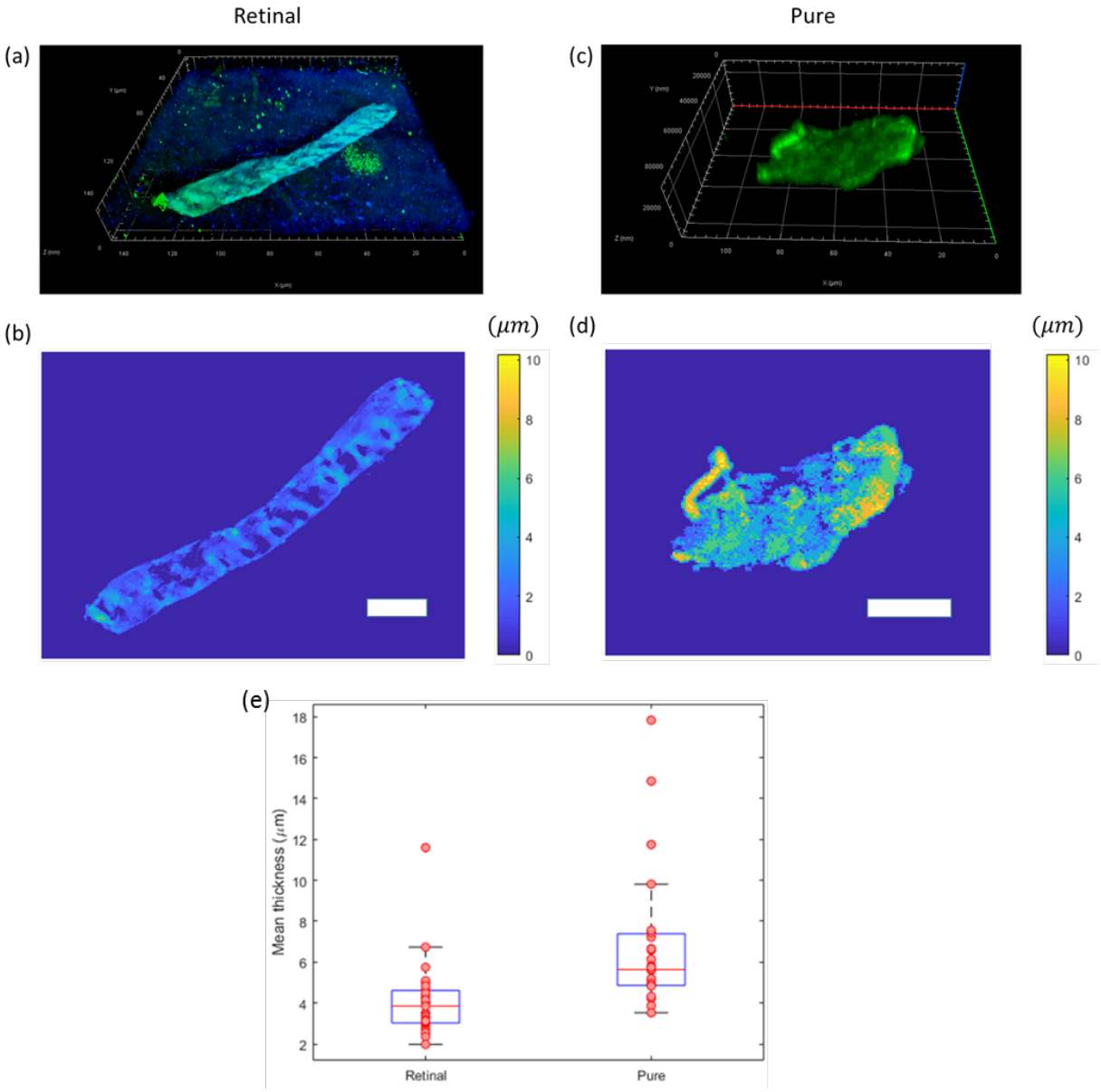


Figure 2.3: (a) 3D view of the 405 nm channel (blue) and 488 nm channel (green). Green represents the Thioflavin signal while blue represents the DAPI signal. (b) Reconstructed thickness image from (a). (c) 3D view of a pure deposit stained with Thioflavin-S. (d) Reconstructed thickness image from (c). Scale bar: 20 μm (e) Boxplot of average thickness of 31 retinal and pure deposits.

The average thickness values of the 31 retinal and 31 pure deposits are plotted in Fig.2.3.(e). The pure deposits (mean ranges from 3.5 μm to 17.8 μm) are generally thicker than the retinal deposits (mean ranges from 2.0 μm to 11.6 μm). Using an unpaired t test, the mean thicknesses of 31 retinal deposits differ significantly from the 31 pure deposits ($p< 10^{-3}$). Hence, for the rest of the analysis in this chapter, only pure deposits with mean thickness lower than 11.60 μm will be used to match the thicknesses of the retinal deposits. This reduces the number of pure deposits from 31 to 28. The mean thickness of the 28 pure deposits are closer to the mean of the 31 retinal deposits but still show a significant difference ($p< 10^{-3}$).
For sample retinal and pure deposits, the registration of thickness and linear retardance are displayed in Fig.2.4. The boundary of the linear retardance and thickness generally matches after registration with minor differences. The overlapping areas in the two maps are shown in yellow in Fig.2.4. This is the area over which birefringence was calculated and analyzed.

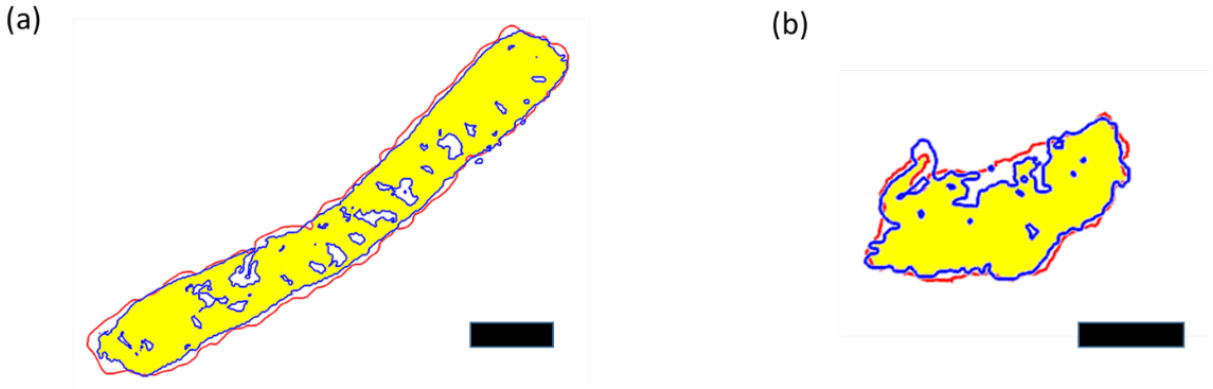


Figure 2.4: Comparison of the boundaries of the linear retardance and thickness after registration of (a) a retinal deposit and (b) a pure deposit. Red contours in both images are the boundaries of linear retardance. Blue contours in both images are the boundaries of thickness. Yellow represents the overlap area of edges of the deposit defined by thickness and retardance. Scale bar: 20 μm.

2.3.3 Polarization properties

Parameters generated from polar decomposition reveal different polarization properties between amyloid deposits and the surrounding retina and between pure Aβ deposits and the surrounding glass. Sample images are shown in Fig.2.5.(a). The average and contrast

of diattenuation, depolarization power and linear retardance are given in Fig.2.5.(b) and Table.2.1. The linear retardance (Fig.2.5.(c), Table.2.1) has the largest relative difference between the deposit and the background. Mean DP and LR of the pure deposits are higher than the retinal deposits while the retina had higher values than the glass. Hence, the contrast in DP and LR are higher in pure deposits than retinal deposits. For 31 retinal deposits, an unpaired t-test shows significant difference in the mean of LR between deposits and surrounding retina ($p< 10^{-9}$). Although no significant differences are found in D ($p= 0.79$) and DP ($p= 0.09$) between the mean of the deposits and retina, the contrast between retina and surround for these properties was significantly above zero (D($p< 10^{-5}$), DP($p< 10^{-7}$)), consistent with the visibility of the deposits. For 28 pure deposits, their mean in LR, D and DP all show significant difference compared to background glass slides (LR: $p < 10^{-5}$, D: $p= 0.002$, DP: $p< 10^{-7}$). Comparing the mean of LR and D of 31 retinal deposits to 28 pure deposits, no significant difference is found using unpaired t test (LR: $p=0.13$, D: $p=0.73$). There is a significant difference in DP between the retinal deposits and pure deposits ($p < 10^{-4}$).

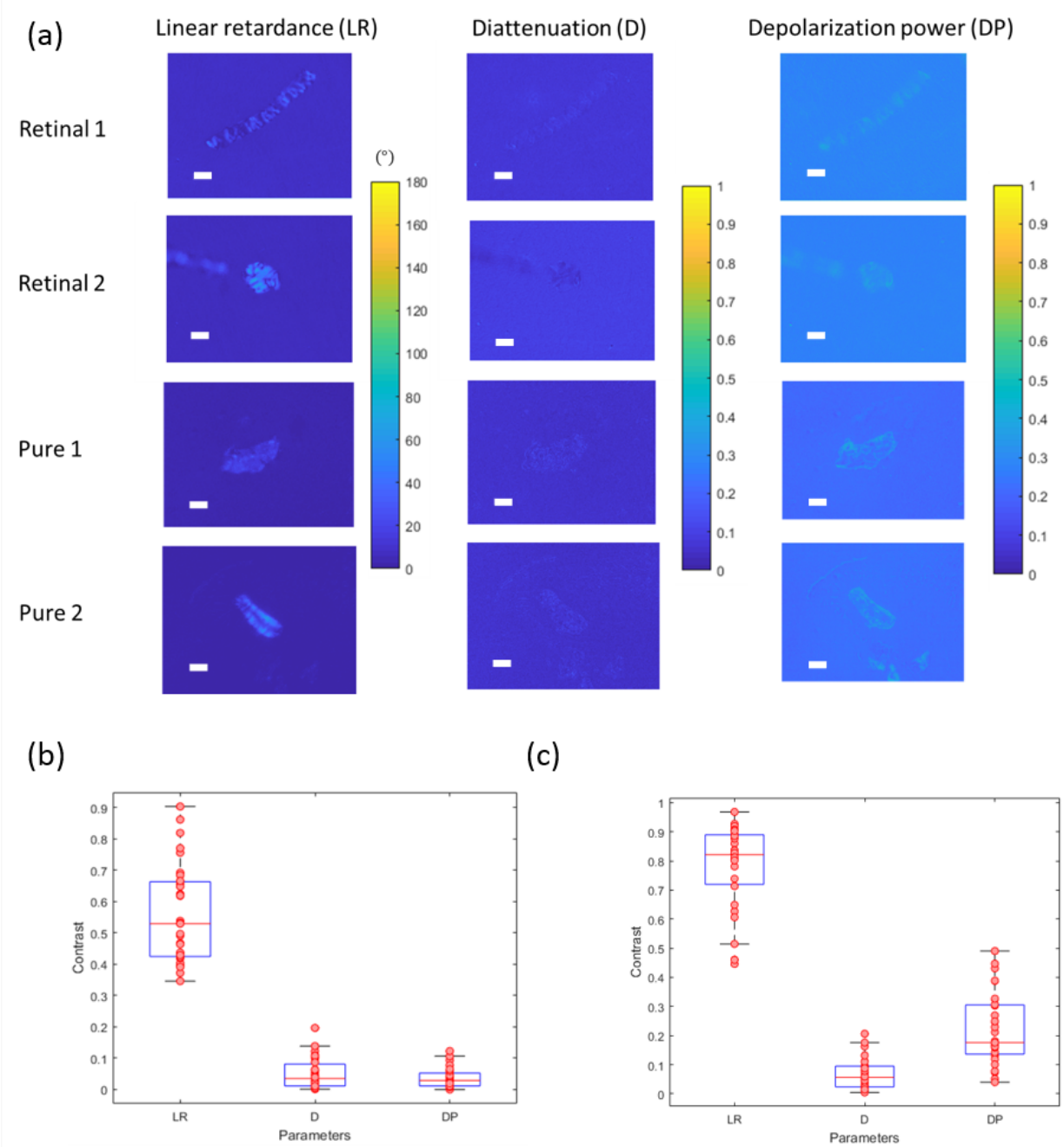


Figure 2.5: (a) Images of LR, D, DP in four deposits and the surrounding retina. The parameters are displayed in full range which LR is [0°, 180°], D is [0°, 1°] and DP is [0, 1]. Scale bar: 20 μm. (b) Boxplot of the contrast of the three parameters of 31 retinal deposits. (c) Box plot of the contrast of the same parameters of 28 pure deposits.

Sample	Mean of	LR (°)	D (0 to 1)	DP (0 to 1)
Retinal	Deposit	24 ± 1	0.070 ± 0.002	0.26 ± 0.01
	Retina	6.6 ± 0.5	0.067 ± 0.003	0.25 ± 0.01
	Contrast	0.56 ± 0.03	0.051 ± 0.009	0.038 ± 0.006
Pure	Deposit	29 ± 4	0.067 ± 0.002	0.34 ± 0.02
	Glass	2.3 ± 0.1	0.060 ± 0.001	0.22 ± 0.01
	Contrast	0.77 ± 0.03	0.063 ± 0.010	0.21 ± 0.02

Table 2.1: Mean and contrast of LR, D, and DP. All values shown are mean and standard error across 31 retinal deposits and 28 pure deposits.

2.3.4 Linear birefringence

Linear retardance, shown in Fig.2.6.(a), increases linearly and significantly with thickness but starts to fluctuate at high thickness giving a fitted linear birefringence of 2.98 ± 0.15 $^\circ/\mu$m. The map of linear birefringence of one retinal deposit, computed pixel by pixel from Eq.2.3 is displayed in Fig.2.6.(b) and has an average value of 9.37 ± 4.14 $^\circ/\mu$m. This average pixel birefringence value is much higher than the fitted linear birefringence. The thickness and fast axis are shown in Fig.2.6.(c), (d). This deposit has an average thickness of 2.5 ± 0.9 μm and most region of the deposit has a fast axis of 140°.

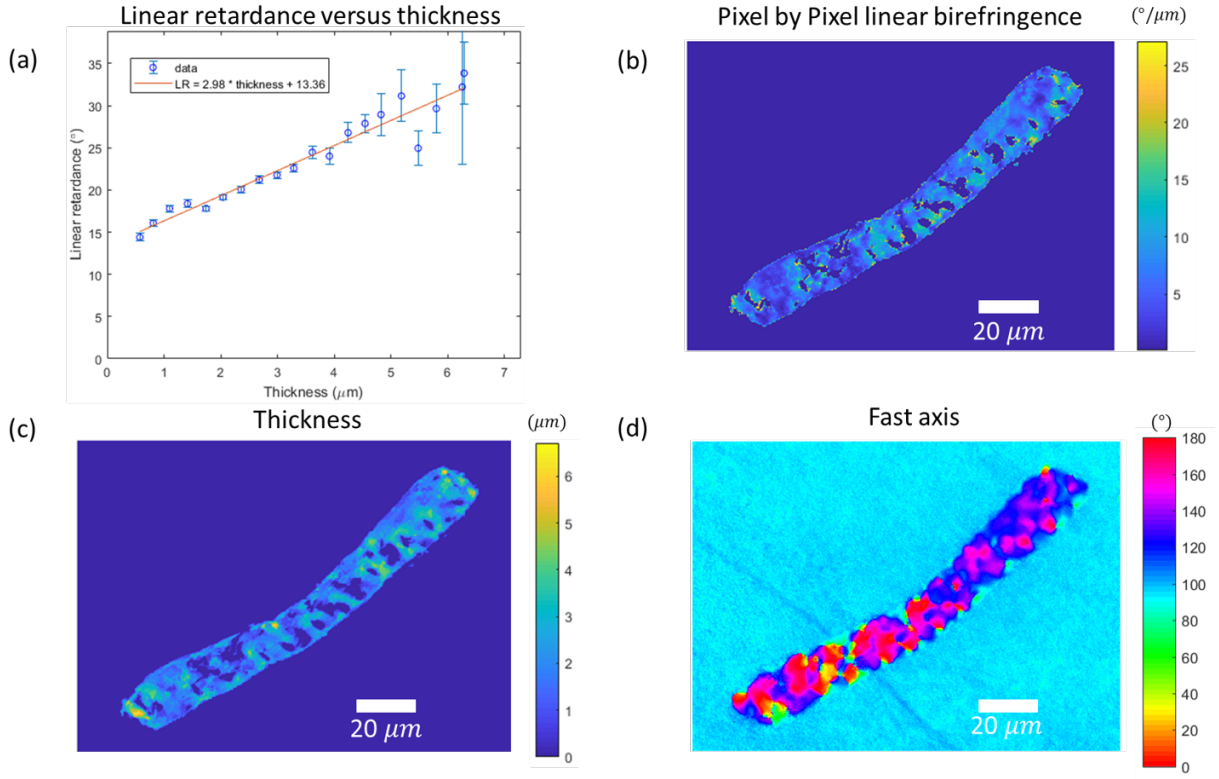


Figure 2.6: (a) Scatter plot and significant linear fit ($p < 10^{-11}$, $R^2 = 0.93$) of linear retardance versus thickness of this retinal deposit. The error bar represents the standard error in linear retardance of pixels with same thickness (b) Pixel by pixel calculation of the linear birefringence of a retinal deposit. (c) Thickness of the deposit. (d) Fast axis of the linear retardance displayed in full range [0°, 180°]. (scale bar = 20 μm)

Across the 31 retinal deposits, 27 showed a significant linear correlation between linear retardance and thickness ($p < 0.05$) across all thickness regions. A retinal deposit with no significant linear correlation between linear retardance across all thickness is shown in Fig.2.7. The linear retardance in higher thickness regions plateaus and begins to gradually decrease for a thickness between 5 and 6 μm. Below a thickness of 5.30 μm, there is a significant linear correlation between the linear retardance and thickness, giving a linear birefringence of 4.56 ± 0.15 °/μm. The average value from a pixel by pixel calculation is 9.45 ± 4.23 °/μm, higher than the fitted linear birefringence. Its thickness and fast axis are shown in Fig.2.7.(c), (d). The mean thickness of this deposit is $3.9 \pm 1.4\mu$m. The lower right region of this deposit has a high thickness (bright yellow region: above 8 μm) as well

as a different fast axis (close to 160°) compared to the rest of the deposit (green region: close to 60°). This region also has a low pixel birefringence close to 2 °/μm, which is lower than both the fitted and pixel birefringence. The fitted birefringence in Fig.2.7.(b) avoids using these high thickness regions.

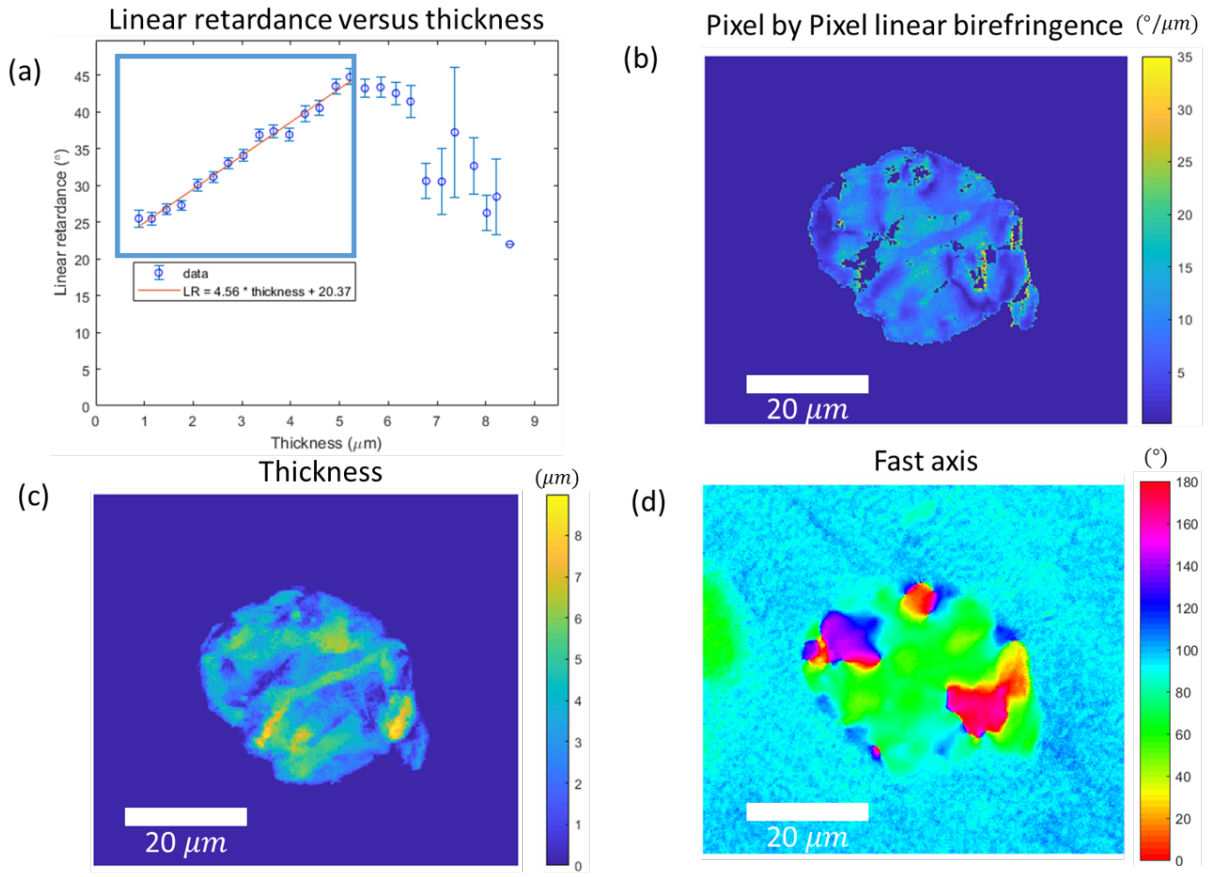


Figure 2.7: (a) Scatter plot and significant linear fit (p< 10^{-12}, R^2 = 0.98) of thickness versus linear retardance of a retinal deposit for a region thinner than 5.30 μm. The error bar represents the standard error in linear retardance of pixels with same thickness. (b) Pixel by pixel calculation of the linear birefringence of a retinal deposit. (c) Thickness of the deposit. (d) Fast axis of the linear retardance displayed in full range [0°, 180°]. (scale bar = 20 μm).

For pure amyloid deposits with thicknesses similar to retinal deposits, only 15 out of 28 deposits show significant linear correlation (p< 0.05) between linear retardance and

thickness across all thickness regions (example Fig.2.8). For the deposit shown in Fig.2.8, the fitted linear birefringence is $2.26 \pm 0.35°/\mu$m. The mean pixel linear birefringence is $6.05 \pm 3.18\ °/\mu$m. Its thickness (mean: $4.9 \pm 1.7\ \mu$m) and fast axis are shown in Fig.2.8.(c),(d). Most regions of the deposit have a thickness close to 6 μm and a fast axis close to 50°. In thicker regions, for example the top and bottom of the deposit in Fig.2.8.(c) (bright yellow: thickness above 8 μm), the fast axis is around 110° (blue regions in the colormap of the fast axis image in Fig.2.8.(d)).

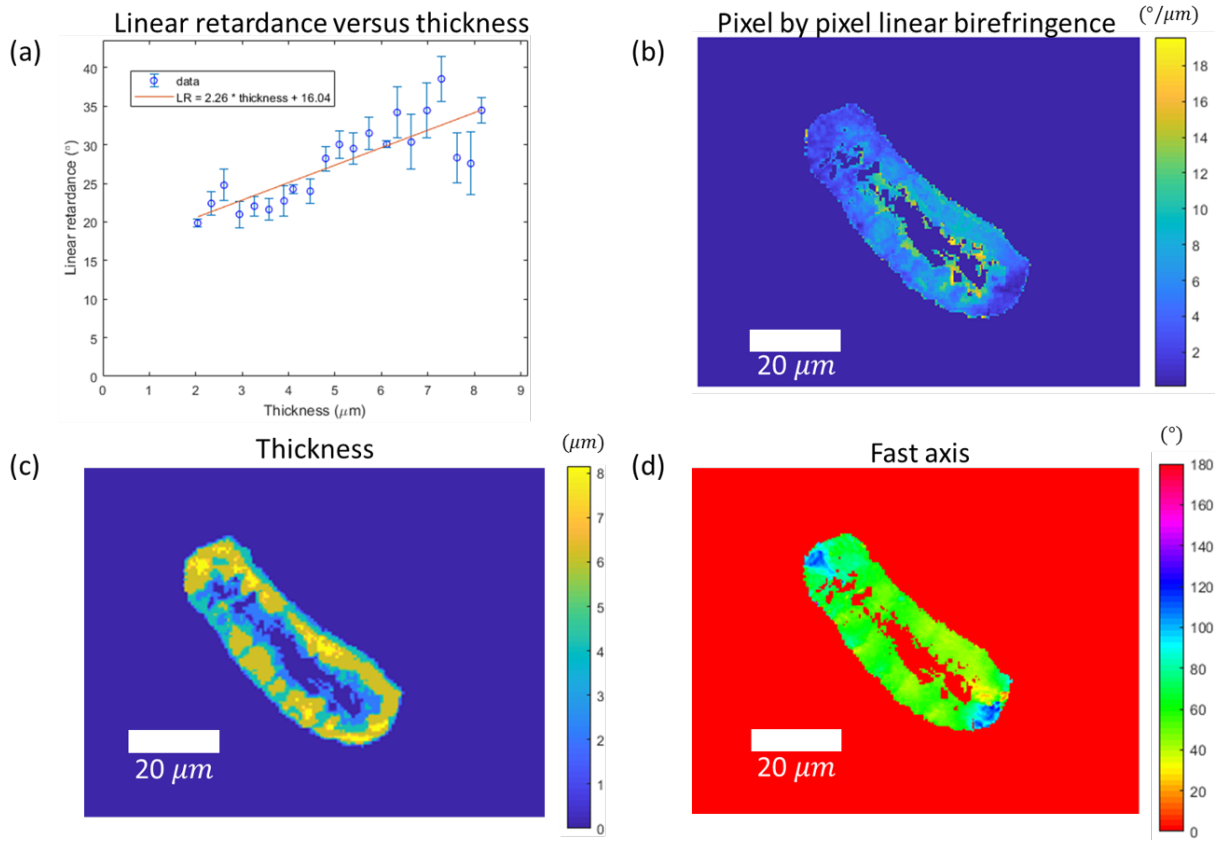


Figure 2.8: (a) Scatter plot and significant linear fit ($p < 10^{-5}$, $R^2 = 0.69$) of thickness versus linear retardance of a pure deposit. A possible area of plateau and decrease occurs beyond 7 μm. The error bars represent the standard error in linear retardance of pixels with same thickness. (b) Pixel by pixel calculation of the linear birefringence. (c) Thickness of the deposits. (d) Fast axis of the linear retardance displayed in full range [0°, 180°]. (scale bar = 20 μm).

A pure deposit with insignificant linear correlation across all thicknesses is displayed in Fig.2.9. The high thickness region also shows a decrease and larger standard errors in linear retardance like the retinal deposit in Fig.2.7. A linear fit to the thickness region lower than 5.80 μm resulted in a fitted linear birefringence of 0.48 ± 0.09 °/μm. The average pixel linear birefringence is 3.21 ± 1.75 °/μm which is also higher than the fitted linear birefringence. The thickness of the deposit has a mean of 4.8 ± 1.8 μm. The high thickness region (bright yellow in Fig.2.9.(c)) generally matches the low pixel birefringence region (dark blue in Fig.2.9.(b)). The fast axis of the tail region on the left of the thickness map also have a different fast axis (close to 10°) than the fast axis of its surrounding (close to 90°).

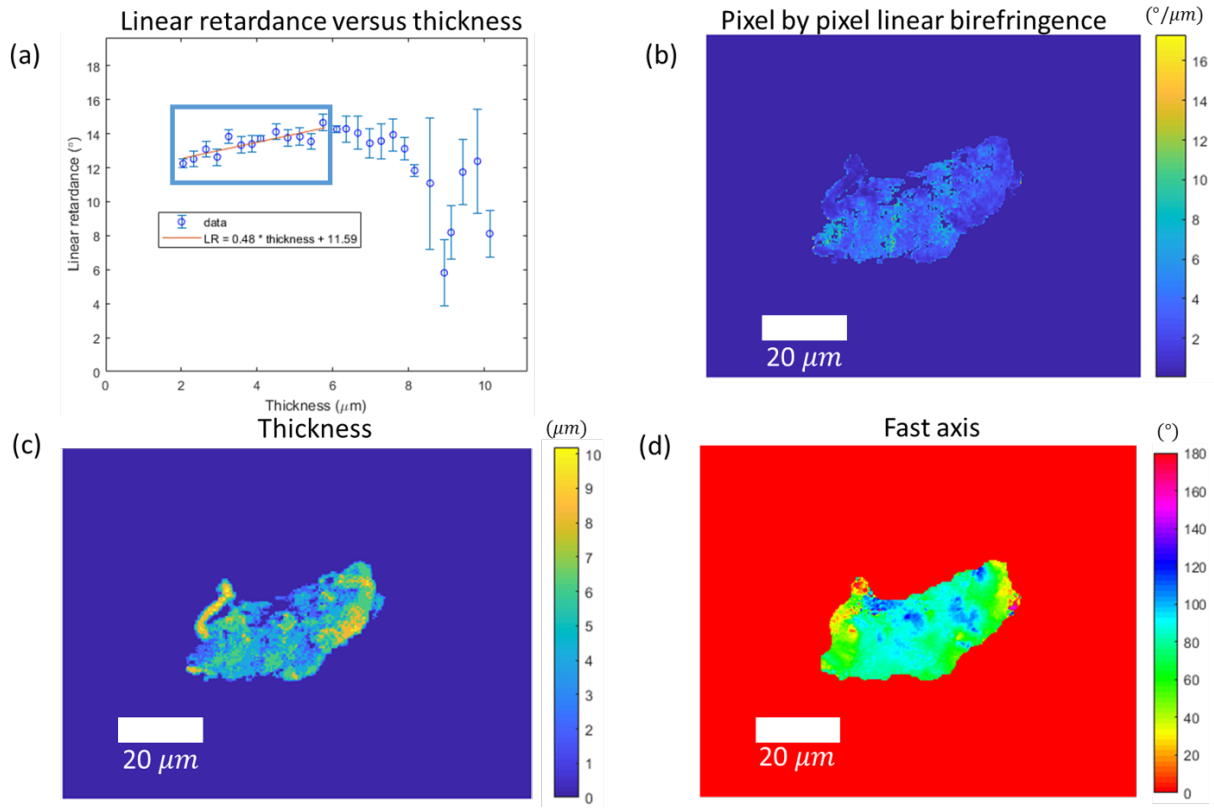


Figure 2.9: (a) Scatter plot and significant linear fit ($p < 10^{-12}$, R^2 = 0.98) of thickness versus linear retardance of a pure deposit for a region thinner than 5.8 μm after which the linear retardance plateaus and then decreases. The error bar represents the standard error in linear retardance of pixels with same thickness. (b) Pixel by pixel calculation of the linear birefringence. (c) Thickness of the deposit. (d) Fast axis of the linear retardance displayed in full range [0°, 180°]. (scale bar = 20 μm).

The linear birefringence calculated from the linear fit of retardance to thickness for lower thickness regions and its pixel by pixel calculation, are compared for the retinal and pure deposits in Fig.2.10.

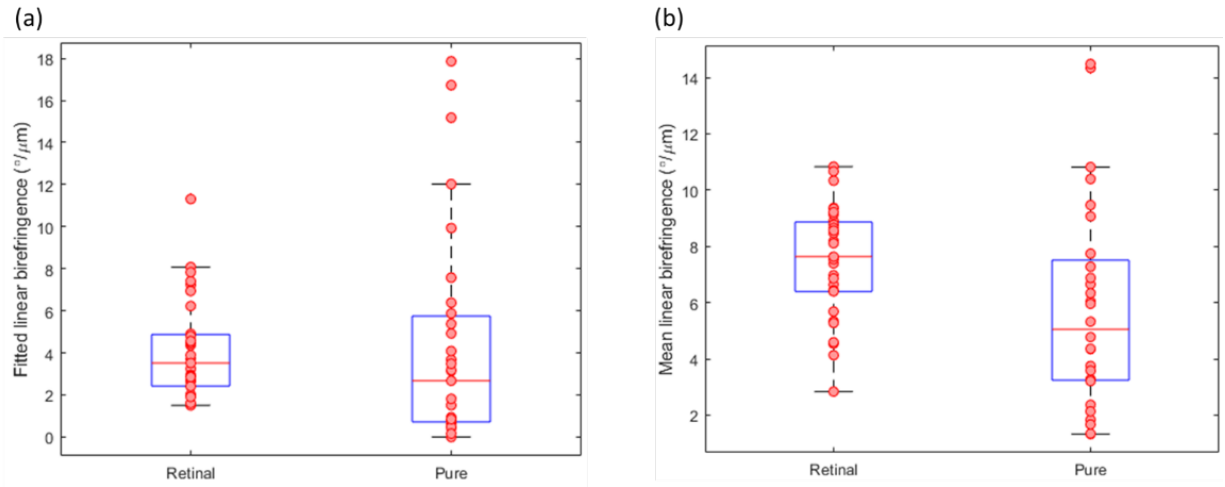


Figure 2.10: (a) Boxplot of fitted linear birefringence value of the retinal and pure deposits. (b) Boxplot of mean pixel by pixel linear birefringence of the retinal and pure deposits.

The linear birefringence of 31 retinal deposits fitted to lower thickness regions, in Fig.2.10.(a), has a mean of 4.18 ± 2.34 °/μm. For 28 pure deposits, the linear birefringence fitted to lower thickness regions, in Fig.2.10.(a), has a mean value of 4.71 ± 5.15 °/μm. An unpaired t test shows no significant difference between the fitted linear birefringence of the retinal and pure deposits ($p = 0.70$). The mean linear birefringence calculated pixel by pixel (Fig.2.10.(b)) of the retinal deposits has a mean of 7.48 ± 1.99 °/μm) and 5.81 ± 3.63 °/μm for pure deposits (Fig.2.10.(b)). An unpaired t test shows a significant difference between the pixel linear birefringence of the retinal and pure deposits ($p = 0.03$). As shown in Fig.2.6, Fig.2.7, Fig.2.8, Fig.2.9, the linear retardance plateaus and then decreases in high thickness regions, indicating a drop in linear birefringence in these regions. Thus, in Fig.2.11. (a), (b), the linear retardance and thickness across all 31 retinal and 28 pure deposits are shown. Each point in the scatter point is an average of linear retardance of all the pixels with the same thickness from different deposits. The scatter plot of retinal deposits (Fig.2.11.(a)) cuts off when the points in the scatter plot only contain pixels from a single deposit (at a thickness of 13.2 μm). To compare with the pure deposits, the scatter plot of the pure deposits is constrained to the same range of thickness. Both plots show descending retardance at high thickness. The significant linear fit of the retinal

deposits ($p < 10^{-37}$, $R^2 = 0.96$) below a thickness of 7.6 μm and pure deposits ($p < 10^{-8}$, $R^2 = 0.80$) below 10.3 μm are plotted. The fitted linear birefringence of these two plots (retinal: 4.44 ± 0.13 °/μm, pure: 4.84 ± 0.29 °/μm) are not significantly different compared to the mean fitted birefringence value (retinal: 4.18 ± 2.34 °/μm, pure: 4.71 ± 5.15 °/μm) given in Fig.2.10.

The effective birefringence, which is the difference in linear retardance divided by the difference in thickness between points in the scatter plot (Fig.2.11.(a),(b)) and the point with the lowest thickness. The birefringence of retinal deposits remains relatively constant until 7 μm, after which the effective birefringence begins to decrease. The effective birefringence of pure deposits fluctuates at a high birefringence value at low thickness. It also starts to decrease around 9 μm.

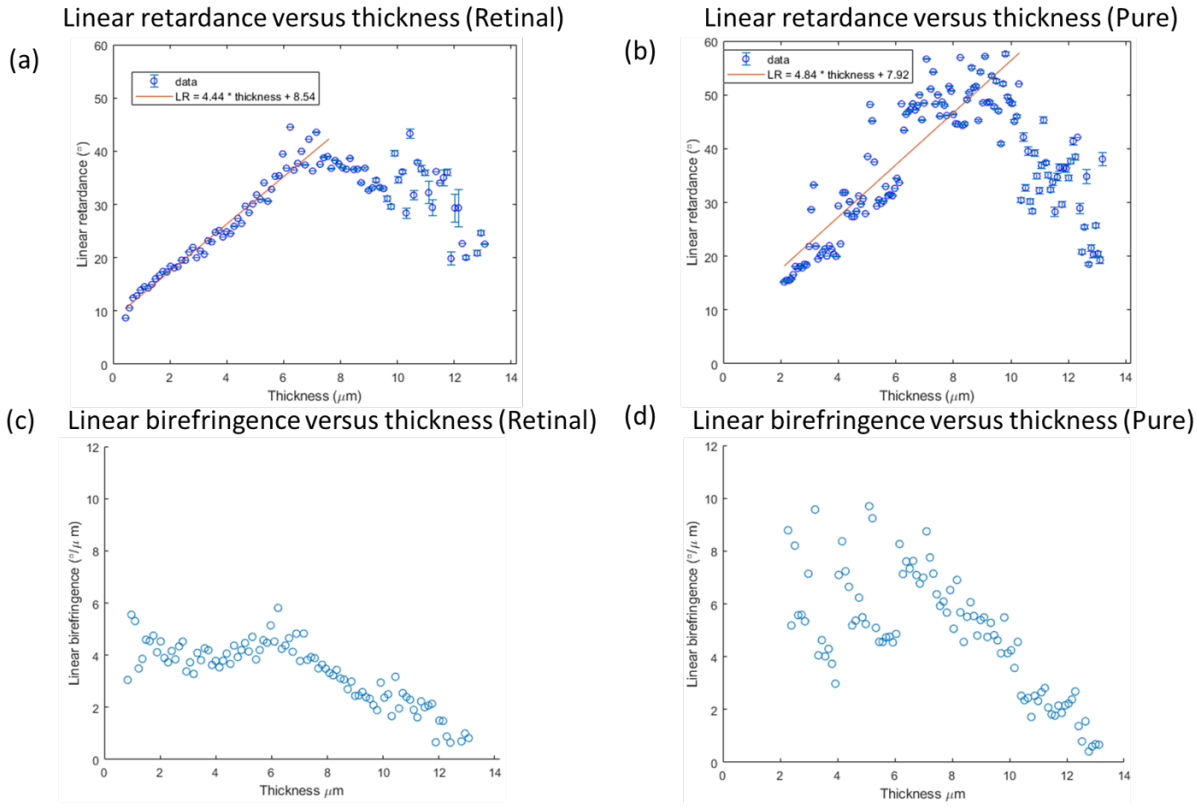


Figure 2.11: Linear retardance versus thickness of (a) retinal and (b) pure deposits. Significant linear fit (retinal: $p < 10^{-37}$, $R^2 = 0.96$, pure: $p < 10^{-24}$, $R^2 = 0.80$) for each plot are shown. The effective linear birefringence versus thickness of (c) retinal, computed from (a), and (d) pure deposits, computed from (b).

2.4 Discussion

Using Mueller Matrix Polarimetry to measure retardance and CSLM to measure thickness combined with image processing, birefringence values of a protein within a tissue can be calculated. However, the accuracy of this method depends on the quality of the image segmentation and registration. Here, image segmentation was accurate enough to approximately match the boundaries in the linear retardance and thickness images.

The measured mean of the single pass linear retardance values of the retina surrounding the deposits in the AD subjects ranges from 4.94° to 11.01°. These values are within the range of single pass linear retardance measured by scanning laser polarimetry (SLP) [66, 50, 79]. Others have suggested that the RNFL in the retina of AD patients is thinner [80, 81, 14] than that of age matched normal and thus would be expected to have lower linear retardance values.

The amyloid deposits are significantly higher in their linear retardance than the surrounding retina. The linear retardance also provides the greatest contrast between them. It is possible that a combination of properties would improve the differentiation of deposit and surround [65, 82].

The average thickness of the individual retinal deposits (ranges from 1.98 μm to 11.60 μm) are within the range of the thickness of the retinal deposits reported by Koronyo *et al.* [20]. The linear retardance tends to drop in higher thickness regions which leads to lower linear birefringence at higher thickness. Since the pure deposits are much thicker in average than the retinal deposits, this result of significant difference in pixel birefringence in retinal and pure deposits could be influenced by the high thickness region. The fitted linear birefringence are computed from linear fit of the linear retardance and thickness at low thickness region. This leads to no significant difference in fitted birefringence between retinal and pure deposits.

The initially increasing retardance at lower thicknesses followed by a decreasing linear retardance at high thickness could be explained by a short range order in the arrangement of fibrils as they grow combined with a lack of long range correlation in the orientation of amyloid fibrils. Other fibrillary structures in the eye also show similar short range order such as crystalline lens fibres [83, 84]. The high intercept of the linear fit in the linear retardance versus thickness plot may also result from a more ordered alignment of fibrils in lower layers in the deposit. The retardance may rise rapidly in the first micron of thickness.

Retinal deposits and pure Aβ deposits on glass of similar thickness show similar retardance and birefringence values that are not significantly different. In both, the retardance plateaus and begins to reduce at higher thicknesses. Thus it may be that the retinal de-

posits being stained by thioflavin in the retina in association with Alzheimers disease are composed predominantly of Aβ, and these fibrillary retinal deposits appear similar to the Aβ cores of neuritic plaques (without the surrounding neuritic material) rather than to diffuse plaques which are usually Thioflavin negative.
Senile plaques, which are presumed to contain Aβ, have been measured by PSOCT in the brain [29]. PSOCT computes the linear birefringence by measuring the depth resolved linear retardance and represents the birefringence as the gradient of the linear retardance [50], similar to the method here which deduces birefringence from a linear fit of linear retardance versus thickness. However, it is not clear if the retardance is linear with depth over the full depth of the measured plaques or whether they show longer range disorder similar to retinal plaques. For thickness > 20 μm, said to be the minimum thickness of brain plaques detectable with PSOCT [29], the birefringence of brain plaques is lower than the measured values for the thinner retinal deposits and is suggestive of the behavior of pure Aβ deposits of similar thickness (> 30 μm) to the brain plaques in which birefringence continues to drop with increasing thickness. The lowest fitted linear birefringence of retinal deposits and pure deposits at thicknesses where it can be deduced from a linear fit of retardance to thickness is 1.51 $^\circ/\mu$m and 0.16 $^\circ/\mu$m. The pixel by pixel birefringence of pure deposits measured decreased to 0.22 $^\circ/\mu$m at a thickness of 38.1 μm. Hence, the difference in birefringence between the senile plaques and retinal deposits could result from thicker and less ordered senile plaques which could also reduce their measured retardance and birefringence.
The birefringent properties of both brain and retinal deposits could result from intrinsic birefringence as well as form birefringence of the fiber bundles. Form birefringence [31], is determined by the difference of refractive index of the surrounding tissue and the fibrils and the spacing between the fibrils and their relative orientation. The refractive index of brain tissue (1.35) [29] and retina (1.36) [67] are close. Even if the senile plaques are composed of the same type of amyloid fibrils as retinal amyloid deposits, difference in their birefringence would arise if the retinal deposits were more compact in structure than senile plaques or the known nonamyloid components in the senile plaques reduced their birefringence. In each case, form birefringence would be higher in retinal deposits.
The birefringence values measured in this Chapter for retinal amyloid deposits and pure Aβ deposits are several orders higher than the value given by flow birefringence [85, 86, 87]. This is reasonable because the flow birefringence experiment measures intrinsic and form birefringence of fibrils in solution and the form birefringence depends on the volume fraction of the molecules and the refractive index of the solvent while measurements in this Chapter is carried out in tissue and measured retardance as a function of thickness of a deposit.
The birefringence and retardance values of the deposits reported in this Chapter are much

higher than previously reported values in the retinal nerve fiber layer (below 0.2 °/μm and mean close to 10° in single pass) [50]. These higher values are the basis of a patented method by one of us for finding these amyloid deposits in vivo using label free polarization sensitive imaging techniques [88]. In spite of the falloff in average birefringence in thicker areas of retinal deposits, every thioflavin positive deposit, analyzed here, had a robust retardance signal much larger on average than that of the optic nerve fiber layer, However, the retardance and birefringence signals may be lower in thicker deposits found in the brain.

2.5 Conclusion

In this chapter, a quantitative measurement of polarimetric properties and thickness of presumed Aβ deposits in the retina and pure Aβ deposits using MMP and CLSM, which provided the linear birefringence values for the deposits, is presented. The similarity of linear birefringence in the retinal deposits to that in pure deposits and the similar variation with thickness gives an insight into the degree of short and long rang order within the deposits. It also suggests that the polarimetric properties of the presumed Aβ deposits in the retina found in association with Alzheimers disease are consistent with high Aβ content. The results also suggest that thicker brain deposits have smaller average birefringence than thinner retinal deposits analogous to the change in average birefringence with thickness of pure Aβ deposits. This method could be extended to other protein deposits within tissue if they can be stained and if the protein's linear retardance is larger than that of the surrounding tissue. The retardance and birefringence of retinal amyloid deposits shown in this paper are higher than values reported for RNFL. These higher values form the basis of a patented label free technique [88] for imaging amyloid in the retina as a biomarker of Alzheimers disease.

Chapter 3

Accelerated polar decomposition of Mueller matrix using GPU

3.1 Introduction

The Mueller matrix acquired from polarimetry contains various parameters related to retardance, diattenuation, and depolarization. To acquire these parameters, decomposition of the Mueller matrix is required. Some proposed methods have been introduced in section 1.5.

One method, which is currently the most widely used, is the polar decomposition [37] proposed by Lu and Chipman in 1996. The method deems the Mueller matrix as a consecutive multiplication of the depolarization, retardance, and diattenuation matrices. Using the properties of each polarimetric matrix, the polar decomposition decomposes the three matrices from the Mueller matrix for computing corresponding polarimetric parameters. Since the polar decomposition arbitrarily decomposes the Mueller matrix into 3 independent optical elements, the change in the order of the three elements can result in change in the decomposition algorithm. The influence of the order of the placement of the three matrices was examined in [89]. It stated that the order of decomposition proposed by Lu and Chipman [37] as well as 2 other placements, in which the depolarization matrix is placed on the left side of the diattenuation matrix, would be optimal to use since it would always result in a physical realizable Mueller matrix. Hence, in this chapter, only the order proposed by Lu and Chipman is considered. In such an order, parameters such as diattenuation and polarizance can be computed from the Mueller matrix without performing the decomposition. So, the parameters computed from the polar decomposition

are the depolarization power from the depolarization matrix, and the retardance from the retardance matrix.

However, a full polar decomposition requires 3 matrix multiplications, 3 matrix divisions, and 1 3×3 eigenvalue decomposition. Although all of these tasks are easy to conduct on a 4×4 matrix, a typical imaging Mueller matrix polarimetry requires computation of these parameters at each pixel. A small 100×100 image would result in 10000 independent decompositions. Looping through each pixel and conducting the same decomposition procedure would be time consuming. However, decomposition at different pixels are independent tasks that require an equal number of steps for each decomposition. Thus, setting up the problem in parallel and computing a large number of pixels at the same time is beneficial for accelerating the processing speed. The current day CPU, for example the latest Intel Core i7-8700 processor, enables multiprocessing on 6 cores with 12 threads which means a maximum of 12 parallel executions at the same time. By distributing the tasks in these cores, one can boost the decomposition speed significantly. However, a graphic processing unit (GPU) has far better multiprocessing power. Currently, GPU has been implemented in large scale to enhance gaming performance, image processing, and deep learning. Especially in deep learning, thousands of matrix multiplications are required at every iteration when using back propagation to tune the network. With this large amount of matrix multiplications, using a GPU can significantly reduce the total training time of a model compared to using a CPU. This is because of the different computational architecture of the GPU. A GPU is composed of dozens of streaming multiprocessors (SM), each containing many streaming processors (SP). In Compute Unified Device Architecture (CUDA), each SM can have thousands of threads. Coupled with the number of SMs in one graphics card, the GPU has the ability to run tens of thousands of threads concurrently. Although each core is not comparable in terms of computation power with respect to a CPU core, the large number of cores in the GPU enables it to conduct simple task in large quantities. In decomposing large quantities of 4×4 Mueller matrices, the number of tasks one can process at an instance outweighs the importance of computation speed. Hence, it is ideal to implement a GPU to boost the speed of computing polarimetric parameters from a large number of Mueller matrices.

In this chapter, an accelerated implementation of the polar decomposition is presented. Some steps in the original method were optimized to be less computationally intensive. Additional efforts were made to replace some iterative steps with analytical solutions which can result in better performance in parallel computing. The results shown here demonstrate the power of implementing GPU processing in either large size Mueller matrix images for instant display or an entire database for statistics.

3.2 Method

3.2.1 Polar decomposition

To demonstrate the steps taken to accelerate the computation, it is important to first write out the basic procedure of polar decomposition. Here, a brief walk-through is presented. A Mueller matrix $\mathbf{M}$ can be represented as block form [36]

$$\begin{aligned}\mathbf{M} &= \begin{pmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{pmatrix} \\ &= m_{00} \begin{pmatrix} 1 & \boldsymbol{D}^T \\ \boldsymbol{P} & \mathbf{m} \end{pmatrix}\end{aligned} \tag{3.1}$$

in which $\mathbf{m}$ is the sub 3×3 matrix from element m_{11} to m_{33}. $\boldsymbol{D}^T = (D_H, D_{45}, D_C)$ is the diattenuation vector where the three entries D_H, D_{45}, D_C are the magnitude of horizontal-vertical, 45°-135°, and circular diattenuation. Thus, from the diattenuation vector, one can directly solve for the linear diattenuation D_L and the total magnitude of diattenuation D as

$$D_L = \sqrt{D_H^2 + D_{45}^2} \qquad D = \sqrt{D_L^2 + D_C^2} \tag{3.2}$$

The vector $\boldsymbol{P} = (P_H, P_{45}, P_C)$ in the Mueller matrix is called polarizance vector, which describes the state of polarization of the output polarized light if the input light is not polarized. The components of the polarizance vector are the horizontal-vertical, 45°-135°, and circular polarizance, from which the linear polarizance and total magnitude of polarizance can be computed as

$$P_L = \sqrt{P_H^2 + P_{45}^2} \qquad P = \sqrt{P_L^2 + P_C^2} \tag{3.3}$$

Polar decomposition then starts with writing the Mueller matrix as

$$\mathbf{M} = \mathbf{M}_\Delta \mathbf{M}_R \mathbf{M}_D \tag{3.4}$$

Each of the three matrices can also be written in block form as

$$\mathbf{M}_\Delta = \begin{pmatrix} 1 & \mathbf{0}^T \\ \boldsymbol{P}_\Delta & \mathbf{m}_\Delta \end{pmatrix} \qquad \mathbf{M}_R = \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{m}_R \end{pmatrix} \qquad \mathbf{M}_D = m_{00} \begin{pmatrix} 1 & \boldsymbol{D}^T \\ \boldsymbol{D} & \mathbf{m}_D \end{pmatrix} \tag{3.5}$$

Since the sub matrix $\mathbf{m}_D$ in $\mathbf{M}_D$ can be directly computed using diattenuation vector $\boldsymbol{D}$ and its magnitude D by

$$\mathbf{m}_D = \sqrt{1-D^2}\,\mathbf{I} + \frac{1-\sqrt{1-D^2}}{D^2}\boldsymbol{D}\boldsymbol{D}^T \tag{3.6}$$

the rest of the task is essentially trying to solve for the polarizance vector $\boldsymbol{P}_\Delta$ and the sub matrix $\mathbf{m}_\Delta$ inside the depolarization matrix, and the sub matrix $\mathbf{m}_R$ inside the retardance matrix. All the elements inside the diattenuation matrix are already known, so it can be decomposed from the Mueller matrix immediately by

$$\begin{aligned}\mathbf{M}\mathbf{M}_D^{-1} &= \mathbf{M}_\Delta\mathbf{M}_R \\ &= \mathbf{M}^{'}\end{aligned} \tag{3.7}$$

where the matrix $\mathbf{M}'$ is the matrix after decomposing the diattenuation matrix from the Mueller matrix. Since the multiplication of the depolarization matrix and retardance matrix is

$$\begin{aligned}\mathbf{M}_\Delta\mathbf{M}_R &= \begin{pmatrix}1 & \mathbf{0}^T\\ \boldsymbol{P}_\Delta & \mathbf{m}_\Delta\end{pmatrix}\begin{pmatrix}1 & \mathbf{0}^T\\ \mathbf{0} & \mathbf{m}_R\end{pmatrix} \\ &= \begin{pmatrix}1 & \mathbf{0}^T\\ \boldsymbol{P}_\Delta & \mathbf{m}_\Delta\mathbf{m}_R\end{pmatrix} \\ &= \mathbf{M}' \\ &= \begin{pmatrix}1 & \mathbf{0}^T\\ \boldsymbol{P}_\Delta & \mathbf{m}'\end{pmatrix}\end{aligned} \tag{3.8}$$

the polarizance vector of the depolarization matrix $\boldsymbol{P}_\Delta$ and multiplication of the two sub matrices can be extracted from the elements of $\mathbf{M}'$. The sub matrix $\mathbf{m}'$, which is essentially $\mathbf{m}_\Delta\mathbf{m}_R$, can be used to extract $\mathbf{m}_\Delta$ by manipulating the unitary matrix property of the sub retardance matrix $\mathbf{m}_R$ (which means $\mathbf{m}_R^T = \mathbf{m}_R^{-1}$) and the symmetric property of the sub depolarization matrix $\mathbf{m}_\Delta$ (which means $\mathbf{m}_\Delta = \mathbf{m}_\Delta^T$). First, it is easy to show that

$$\begin{aligned}\mathbf{m}'\mathbf{m}'^T &= \mathbf{m}_\Delta\mathbf{m}_R\mathbf{m}_R^T\mathbf{m}_\Delta^T \\ &= \mathbf{m}_\Delta\mathbf{m}_R\mathbf{m}_R^{-1}\mathbf{m}_\Delta \\ &= \mathbf{m}_\Delta^2\end{aligned} \tag{3.9}$$

Then using Cayley-Hamilton theory, which dictates that a square matrix satisfies its own characteristic polynomial, plugging $\mathbf{m}_\Delta$ into its own characteristic polynomial yields (See Appendix A)

$$\begin{aligned}\mathbf{m}_\Delta &= [\mathbf{m}_\Delta^2 + (\lambda_1'\lambda_2' + \lambda_2'\lambda_3' + \lambda_3'\lambda_1')\mathbf{I}]^{-1}[(\lambda_1' + \lambda_2' + \lambda_3')\mathbf{m}_\Delta^2 + \lambda_1'\lambda_2'\lambda_3'\mathbf{I}] \\ &= [\mathbf{m}'\mathbf{m}'^T + (\lambda_1'\lambda_2' + \lambda_2'\lambda_3' + \lambda_3'\lambda_1')\mathbf{I}]^{-1}[(\lambda_1' + \lambda_2' + \lambda_3')\mathbf{m}'\mathbf{m}'^T + \lambda_1'\lambda_2'\lambda_3'\mathbf{I}]\end{aligned} \tag{3.10}$$

where $(\lambda_1', \lambda_2', \lambda_3')$ are the eigenvalues of $\mathbf{m}_\Delta$, which are equivalent to the square root of the eigenvalues of $\mathbf{m}'\mathbf{m}'^T$ (See eq.3.9) which is represented by $(\lambda_1, \lambda_2, \lambda_3)$. Then Eq.3.10 becomes

$$\mathbf{m}_\Delta = \pm[\mathbf{m}'\mathbf{m}'^T + (\sqrt{\lambda_1\lambda_2} + \sqrt{\lambda_2\lambda_3} + \sqrt{\lambda_3\lambda_1})\mathbf{I}]^{-1}[(\sqrt{\lambda_1} + \sqrt{\lambda_2} + \sqrt{\lambda_3})\mathbf{m}'\mathbf{m}'^T + \sqrt{\lambda_1\lambda_2\lambda_3}\mathbf{I}] \tag{3.11}$$

where $\pm$ is determined by the determinant of $\mathbf{m}'$. The depolarization power is then computed from the depolarization matrix by

$$\Delta = 1 - \frac{|\mathrm{tr}(\mathbf{m}_\Delta)|}{3} \tag{3.12}$$

Then $\mathbf{m}_R$ can be easily decomposed as $\mathbf{m}_\Delta^{-1}\mathbf{m}'$. The sub retardance matrix is equivalent to a 3×3 rotation matrix, from which the retardance R and rotation axis $\boldsymbol{R}$ (retardance vector) can be extracted using

$$R = \cos^{-1}\left[\frac{\mathrm{tr}(\mathbf{m}_R) - 1}{2}\right] \tag{3.13}$$

and

$$\boldsymbol{R} = \frac{1}{2\sin R}\begin{pmatrix}(\mathbf{m}_R)_{12} - (\mathbf{m}_R)_{21}\\ (\mathbf{m}_R)_{20} - (\mathbf{m}_R)_{02}\\ (\mathbf{m}_R)_{01} - (\mathbf{m}_R)_{10}\end{pmatrix} \tag{3.14}$$

The components of the retardance vector $\boldsymbol{R}$ are the horizontal-vertical retardance R_H, 45°-135° retardance R_{45}, and circular retardance R_C. The linear retardance R_L is then computed as

$$R_L = \sqrt{(R_H)^2 + (R_{45})^2} \tag{3.15}$$

This concludes the polar decomposition with all the parameters, including $\boldsymbol{D}$, $\boldsymbol{P}$, Δ, $\boldsymbol{P}_\Delta$, and $\boldsymbol{R}$.

3.2.2 Accelerated implementation

The above process gives the original procedure for polar decomposition of the Mueller matrix. However, the process involves computations on 4×4 matrices, which are much more computationally intensive than computation on 3×3. In this section, an accelerated implementation is used to conduct the whole process on 3×3 matrices.

First, Eq.3.4 can be written as

$$\begin{aligned}\mathbf{M} &= \begin{pmatrix} 1 & \mathbf{0}^T \\ \boldsymbol{P}_\Delta & \mathbf{m}_\Delta \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & \mathbf{m}_R \end{pmatrix} \begin{pmatrix} 1 & \boldsymbol{D}^T \\ \boldsymbol{D} & \mathbf{m}_D \end{pmatrix} \\ &= \begin{pmatrix} 1 & \boldsymbol{D}^T \\ \boldsymbol{P}_\Delta + \mathbf{m}_\Delta \mathbf{m}_R \boldsymbol{D} & \boldsymbol{P}_\Delta \boldsymbol{D}^T + \mathbf{m}_\Delta \mathbf{m}_R \mathbf{m}_D \end{pmatrix}\end{aligned} \tag{3.16}$$

Hence, matching the elements with 3.4, it is easy to get

$$\boldsymbol{P}_\Delta = \frac{\boldsymbol{P} - \mathbf{m}\boldsymbol{D}}{1 - D^2} \tag{3.17}$$

and

$$\begin{aligned}\mathbf{m}_\Delta \mathbf{m}_R &= (\mathbf{m} - \boldsymbol{P}_\Delta \boldsymbol{D}^T)\mathbf{m}_D^{-1} \\ &= \mathbf{m}'\end{aligned} \tag{3.18}$$

Once $\boldsymbol{P}_\Delta$ is computed, it is clear that $\mathbf{m}'$ can be directly solved by getting the inverse of $\mathbf{m}_D$. Since $\boldsymbol{D}$ satisfies

$$\boldsymbol{D}^T \boldsymbol{D} = D^2 \tag{3.19}$$

the inverse of $\mathbf{m}_D$ can be written in the form of $\alpha \boldsymbol{I} + \beta \boldsymbol{D}\boldsymbol{D}^T$ which matches with the expression of $\mathbf{m}_D$ in Eq.3.6 so that the term $\boldsymbol{D}\boldsymbol{D}^T\boldsymbol{D}\boldsymbol{D}^T$ can be reduced to $D^2\boldsymbol{D}\boldsymbol{D}^T$. Solving α and β, the inverse of $\mathbf{m}_D$ can be directly written as

$$\mathbf{m}_D^{-1} = \frac{\sqrt{1 - D^2} - 1}{\sqrt{1 - D^2} D^2} \boldsymbol{D}\boldsymbol{D}^T + \frac{1}{\sqrt{1 - D^2}} \mathbf{I} \tag{3.20}$$

Hence, instead of computing the full 4×4 diattenuation matrix $\mathbf{M}_D$ and then inverting it, one could directly compute the inverse of the $\mathbf{m}_D$ and use it to decompose $\mathbf{m}'$. This manipulation reduces the original 4×4 matrix division to 3×3 matrix multiplication.
Then, the problem is reduced to solving $\mathbf{m}_\Delta$ using Eq.3.11 which requires computing the eigenvalues of $\mathbf{m}'(\mathbf{m}')^T$. For computing the eigenvalues of a matrix, iterative approaches are used in most cases. However, in this case, a faster approach can be adopted. Since the product of a matrix with its transpose is a symmetric matrix, $\mathbf{m}'(\mathbf{m}')^T$ is a symmetric matrix. For a real symmetric 3×3 matrix, one can compute its eigenvalues analytically. This analytical solution is the largest contributor for improving speed in the polar decomposition. As shown in [90], in order to express the eigenvalues, three auxiliary parameters b, p, and q can be created using the elements in $\mathbf{m}'$ as

$$b = \mathbf{m}'_{00} + \mathbf{m}'_{11} + \mathbf{m}'_{22} \tag{3.21}$$

$$p = \frac{1}{2}[(\mathbf{m}'_{00} - \mathbf{m}'_{11})^2 + (\mathbf{m}'_{00} - \mathbf{m}'_{22})^2 + (\mathbf{m}'_{11} - \mathbf{m}'_{22})^2] + 3((\mathbf{m}'_{01})^2 + (\mathbf{m}'_{02})^2 + (\mathbf{m}'_{12})^2) \quad (3.22)$$

$$\begin{aligned} q =& 18(\mathbf{m}'_{00}\mathbf{m}'_{11}\mathbf{m}'_{22} + 3\mathbf{m}'_{01}\mathbf{m}'_{02}\mathbf{m}'_{12}) + 2((\mathbf{m}'_{00})^3 + (\mathbf{m}'_{11})^3 + (\mathbf{m}'_{22})^3) + 9(\mathbf{m}'_{00} + \mathbf{m}'_{11} + \\ & \mathbf{m}'_{22})((\mathbf{m}'_{01})^2 + (\mathbf{m}'_{02})^2 + (\mathbf{m}'_{12})^2) - 3(\mathbf{m}'_{00} + \mathbf{m}'_{11})(\mathbf{m}'_{00} + \mathbf{m}'_{22})(\mathbf{m}'_{11} + \mathbf{m}'_{22}) - 27(\\ & \mathbf{m}'_{00}(\mathbf{m}'_{12})^2 + \mathbf{m}'_{11}(\mathbf{m}'_{02})^2 + \mathbf{m}'_{22}(\mathbf{m}'_{01})^2) \end{aligned} \quad (3.23)$$

Using these three auxiliary parameters, the three eigenvalues $\lambda_1, \lambda_2, \lambda_3$ in Eq.3.11 are computed as

$$\lambda_1 = \frac{1}{3}\Big[b + 2\sqrt{p}\cos\Big(\frac{1}{3} \times \arccos\Big(\frac{q}{2\sqrt{p^3}}\Big)\Big)\Big] \quad (3.24)$$

$$\lambda_2 = \frac{1}{3}\Big[b + 2\sqrt{p}\cos\Big(\frac{1}{3} \times \Big(\arccos\Big(\frac{q}{2\sqrt{p^3}}\Big) + 2\pi\Big)\Big)\Big] \quad (3.25)$$

$$\lambda_2 = \frac{1}{3}\Big[b + 2\sqrt{p}\cos\Big(\frac{1}{3} \times \Big(\arccos\Big(\frac{q}{2\sqrt{p^3}}\Big) - 2\pi\Big)\Big)\Big] \quad (3.26)$$

Thus, using the above expression of the eigenvalues analytically, one can directly compute $\mathbf{m}_\Delta$ and Δ in Eq.3.12. However, computing the sub depolarization matrix for its depolarization power Δ and inverting it for computing $\mathbf{m}_R$ would be redundant. First, the depolarization power is computed from the trace of the sub depolarization matrix $\mathbf{m}_\Delta$ which is the same as the sum of its eigenvalues. Hence, the depolarization power Δ is

$$\Delta = 1 - \frac{\sqrt{\lambda_1} + \sqrt{\lambda_2} + \sqrt{\lambda_3}}{3} \quad (3.27)$$

Second, the sign of the determinant of the sub depolarization matrix is the same as its inverse. Hence, instead of computing the sub depolarization matrix then inverting it, one can directly compute its inverse by inverting Eq.3.11

$$\begin{aligned} (\mathbf{m}_\Delta)^{-1} = \pm\, & [(\sqrt{\lambda_1} + \sqrt{\lambda_2} + \sqrt{\lambda_3})\mathbf{m}'\mathbf{m}'^T + \sqrt{\lambda_1\lambda_2\lambda_3}\mathbf{I}]^{-1} \\ & \times [\mathbf{m}'\mathbf{m}'^T + (\sqrt{\lambda_1\lambda_2} + \sqrt{\lambda_2\lambda_3} + \sqrt{\lambda_3\lambda_1})\mathbf{I}] \end{aligned} \quad (3.28)$$

where the $\pm$ sign is determined by the sign of the determinant of $(\mathbf{m}_\Delta)^{-1}$. This can save one more step for matrix division. Finally, the $\mathbf{m}_R$ is decomposed by multiplying the $(\mathbf{m}_\Delta)^{-1}$ with $\mathbf{m}'$ on the right.

3.2.3 Implementation on the GPU

The accelerated version of the decomposition was implemented in parallel on the GPU (NVIDIA GeForce GTX 1050 (2Gb, 1435MHz)) and the CPU (Intel Core i5-4590 @ 3.30GHz). The result was compared with the original method implemented also in parallel. For parallel computing on the GPU and CPU, the MATLAB Parallel Computing toolbox [91] was used. To implement GPU assisted parallel computing on the GPU, the Mueller matrix images were stored as tensors and processed using batch matrix multiplication and division functions in the toolbox. For CPU parallel computing, 4 threads were created to compute the Mueller matrix stored as a long 3D array. The three methods were benchmarked on runtime versus different sizes of the images, which means different numbers of Mueller matrices fed to the code.
The performance of using a GPU in computing the Mueller matrices was also compared with computing them in parallel on the CPU.

3.3 Results

3.3.1 Results of Mueller matrix computation

The result of performance using GPU and CPU for computing Mueller matrices is shown in Fig.3.1

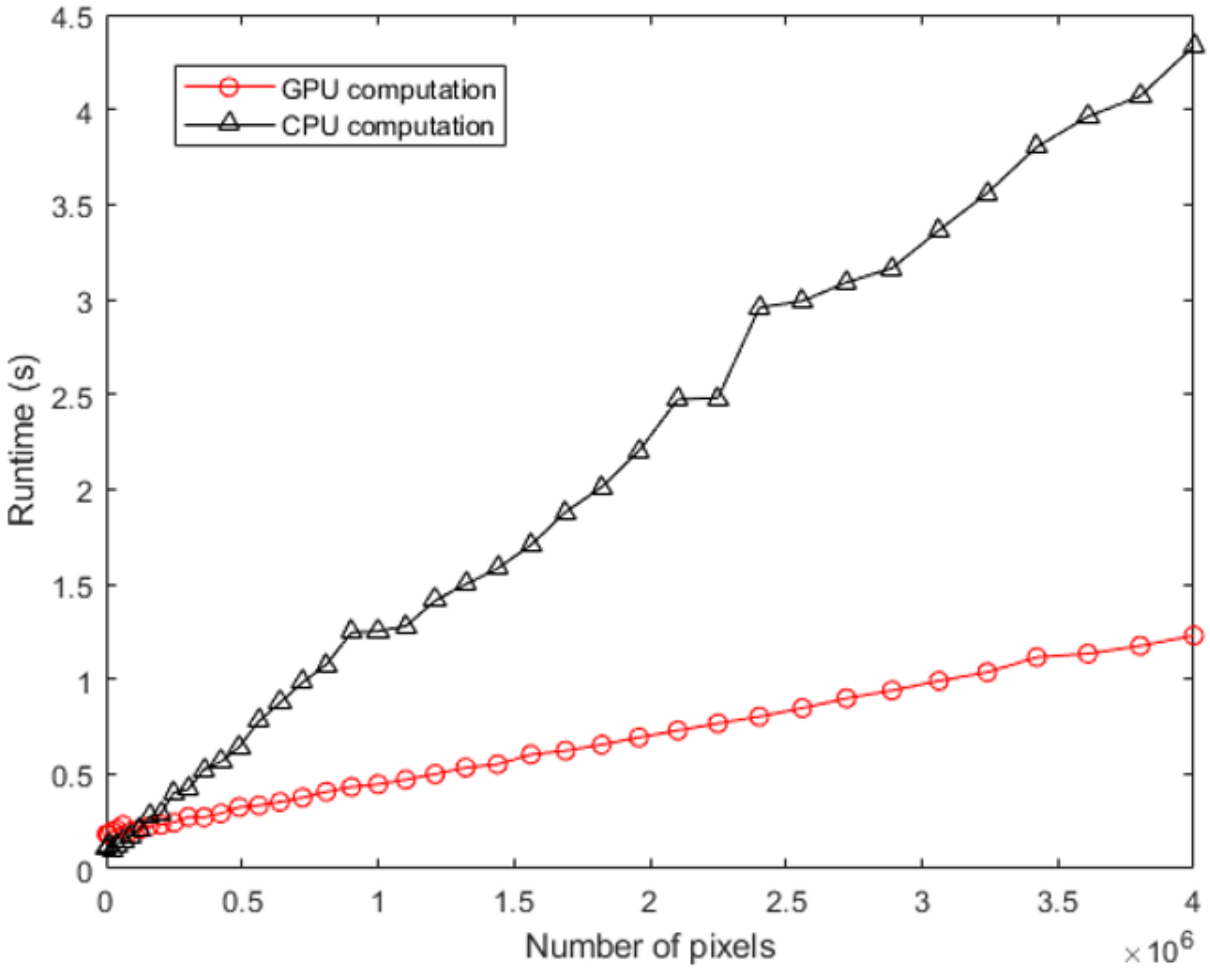


Figure 3.1: Computation time (Runtime) of Mueller matrix compared to number of pixels of the input image

As shown, the GPU outperformed the CPU in computing the Mueller matrix from the 16 images. As image size went up, the speed advantage of the GPU also became larger. Even when 16 intensity images are of sizes 2000×2000, the GPU can still compute the Mueller matrices of all pixels in 1.5s.
From the Mueller matrix, as illustrated in section 1.5, a number of parameters can be directly solved without decomposition. These parameters are depolarization index (DI), anisotropy (linear: A_L, circular: A_C), Q metric (Q), diattenuation (D_L, D_C, angle: D_θ), and polarizance (P_L, P_C, angle: P_θ). In Fig.3.2, the runtime of computing these parameters

(including the time of computing the Mueller matrix in Fig.3.1) is plotted.

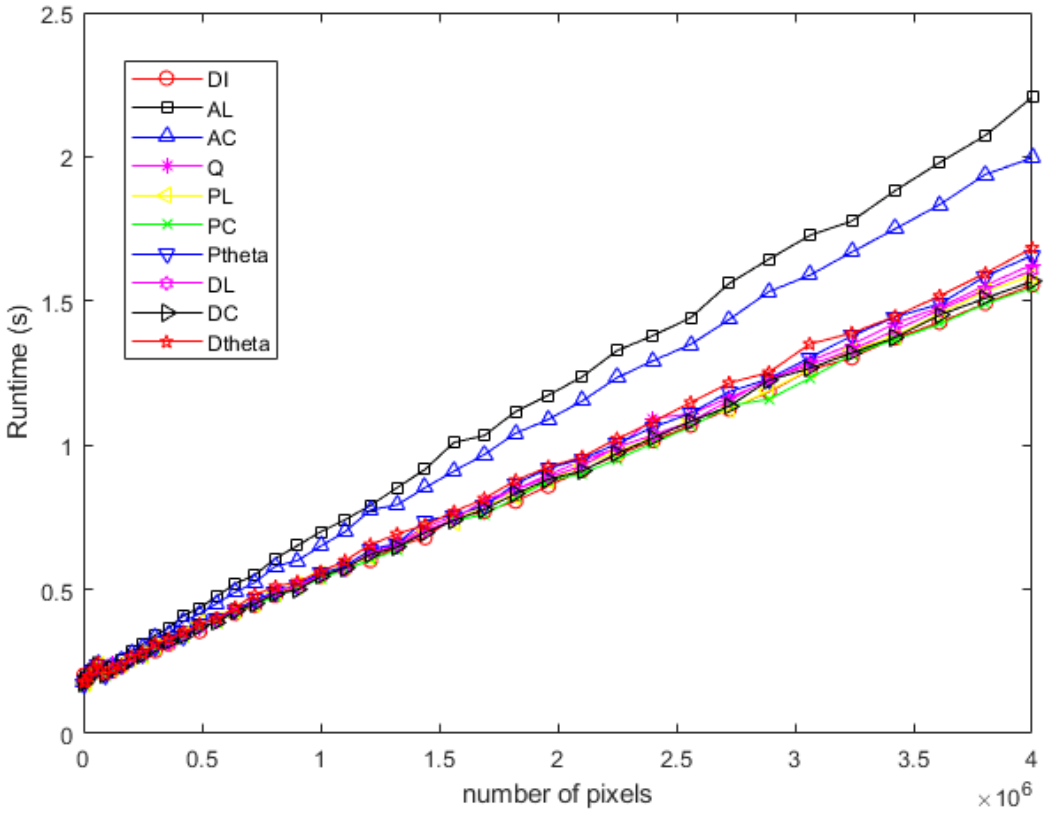


Figure 3.2: Computation time for the Mueller matrix and each of DI, Q, A_L, A_C, Q, D_L, D_C, D_θ, P_L, P_C, P_θ versus the number of pixels of the input image

For up to 4×10^6 pixels, the same number of pixels as a 2000×2000 image, all the parameters except A_L and A_C can be computed within 2 seconds upon acquiring the 16 intensity images. The computation time for A_L and A_C are slightly longer as the computation requires more intermediate steps (shown in section 1.5) than other parameters.

3.3.2 Result of accelerated implementation of polar decomposition

As for polar decomposition, a full set of decomposition presented here includes decomposition of the Mueller matrix to 3 parameters in diattenuation, 3 parameters in polarizance, 3 parameters in retardance, and 1 parameter in depolarization. Strictly speaking, the parameters corresponding to diattenuation and polarizance do not require polar decomposition to solve. However, they are included in the runtime to offer a complete evaluation for the accelerated implementation. First, the runtime of parallel computing of the decomposition on the CPU using the original algorithm was compared to the accelerated algorithm also running on the CPU. The result is shown in Fig.3.3.

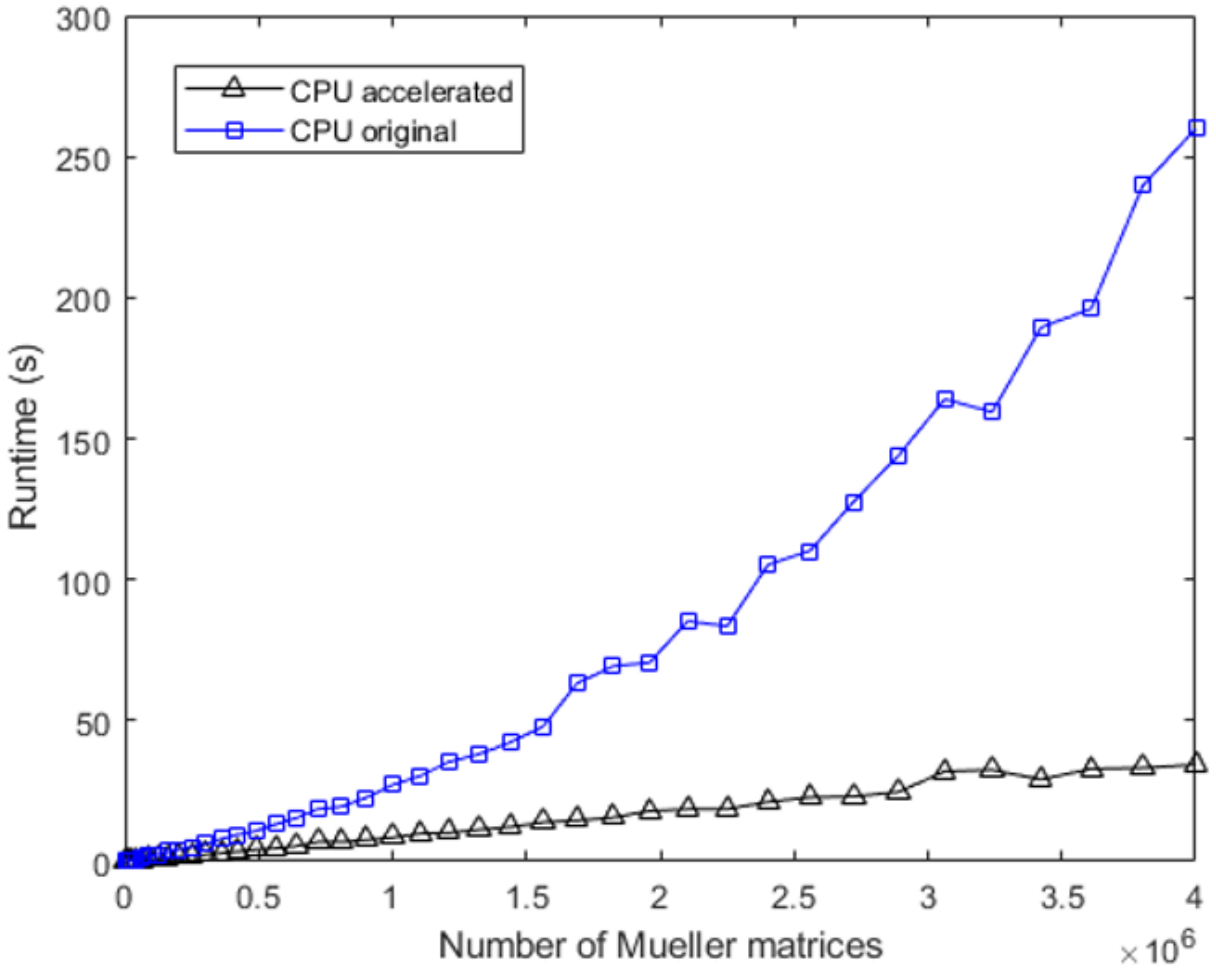


Figure 3.3: Computation time for implementing accelerated polar decomposition and original polar implementation of the Mueller matrix on the CPU versus the number of pixels

It is clear that the optimized method significantly decreased the processing time as the image size became larger and larger. At the image size of 2000×2000, the runtime of the accelerated implementation is more than 5 times faster than the original method. Then, comparing the implementation of the accelerated method on the CPU to GPU, Fig.3.4 on runtime can be plotted as

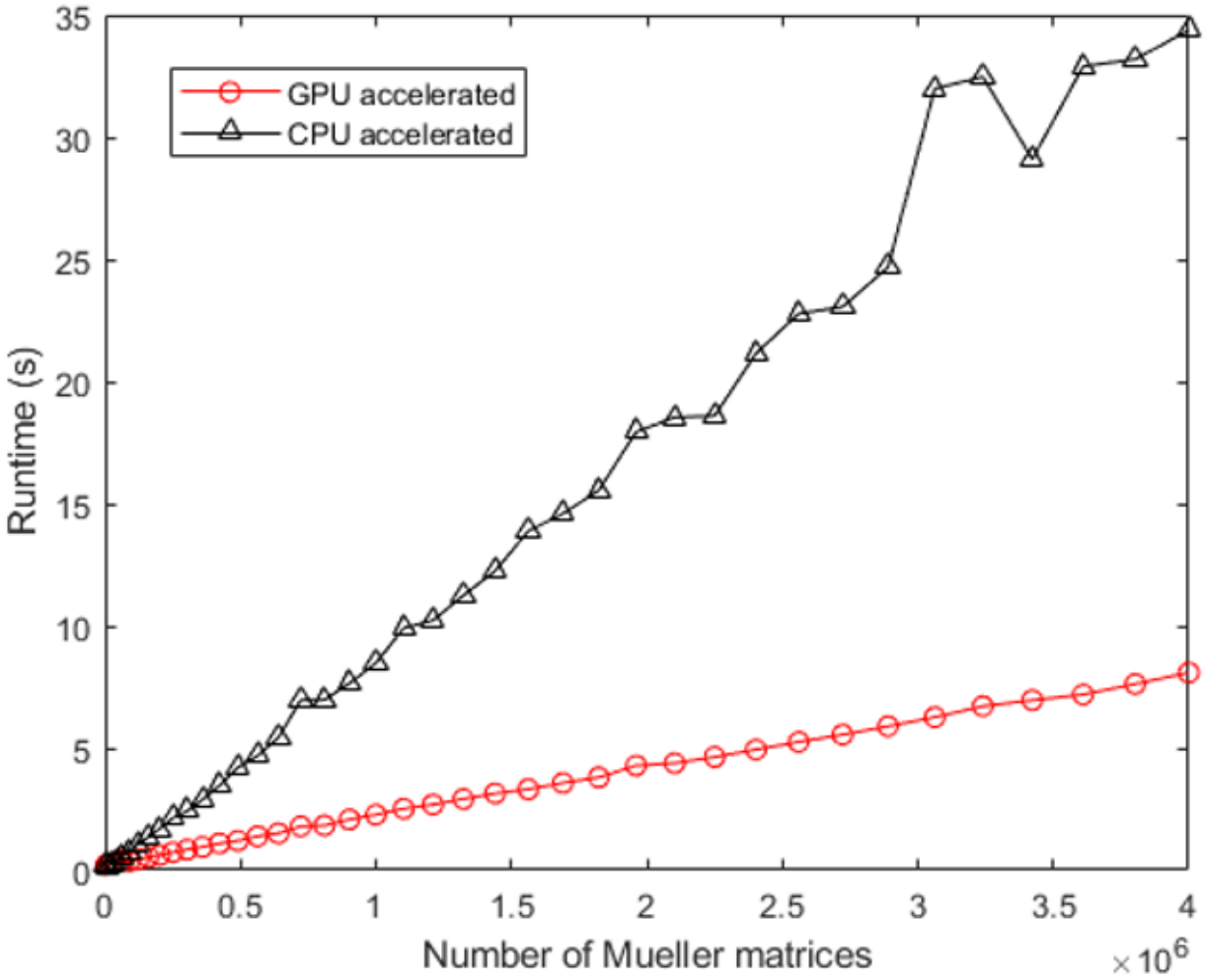


Figure 3.4: Computation time for accelerated polar decomposition of all the Mueller matrix parameters on the CPU and GPU versus the number of pixels

As shown, the GPU empowered the accelerated method to decompose all the parameters of a 2000×2000 Mueller matrix image, calculating them and transferring back to memory within 10s. For a 1000×1000 image, which is close to the image size acquired by the instrument used in this thesis, the runtime is about 2s. This is more than 15 times faster than the original method on the CPU which would take about 30s in Fig.3.3. Since this runtime figure includes all the parameters from the polar decomposition, the runtime would be faster for only extracting specific parameters of interest that require polar decomposition to solve. These parameters are Δ, R_L, R_C, and R_θ. The runtime for solving these 4

parameters separately from the polar decomposition, including the time required from computing Mueller matrix from the 16 intensity images, is plotted in Fig.3.5.

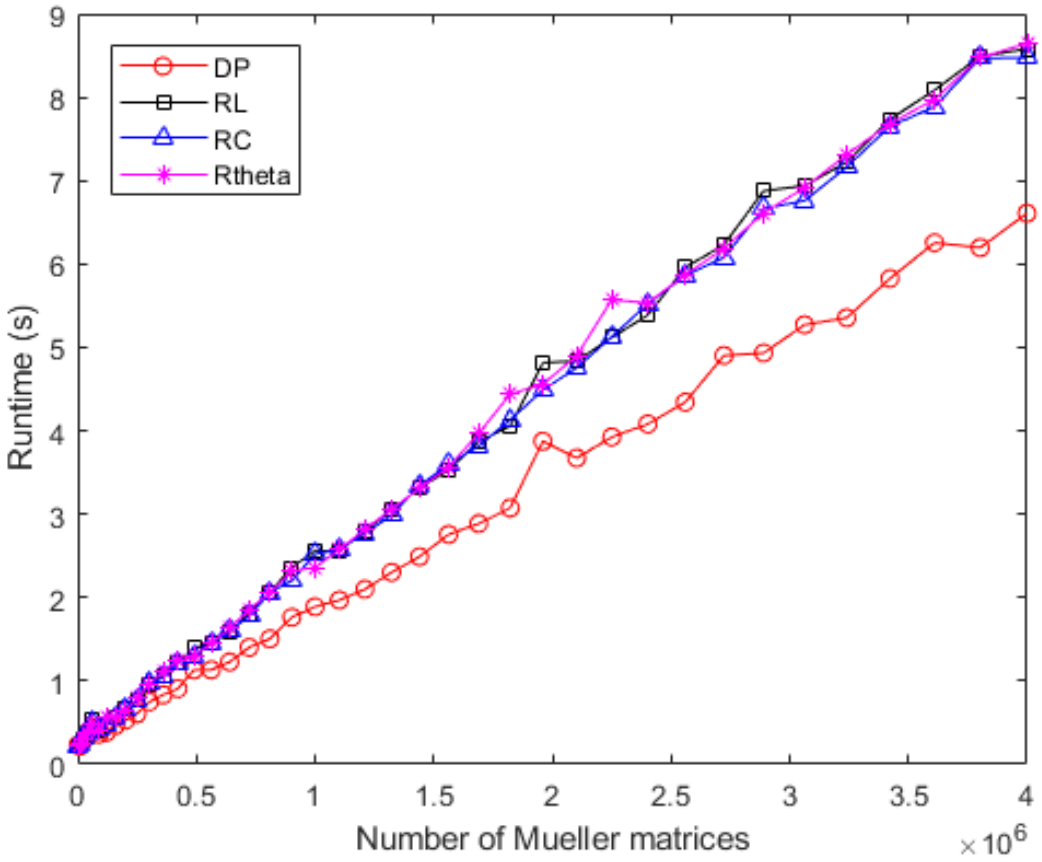


Figure 3.5: Computation time for computing specific parameters, Δ, R_L, R_C, and R_θ, following accelerated polar decomposition on the GPU versus the number of pixels

The computation speed is slower than for the parameters shown in Fig.3.2. Only images smaller than 500×500 can be computed for the linear retardance within 1s.

3.4 Discussion

The result of the accelerated implementation showed significant improvement in computation speed of the polarimetric properties. It indicated that bringing the full decomposition from 4×4 matrix computation to 3×3 computation and using an analytical solution for eigenvalues can accelerate the full process to several times faster, while still using a CPU. Then, using the same accelerated implementation, a GPU helped to bring the process even several times faster than computing on the CPU. In addition to acceleration, using the 3×3 notation greatly reduces the memory occupancy of the large number of matrices on the GPU. This reduced memory requirement for storing the data in the GPU extends the ability of a memory card for processing large Mueller matrix images.

The key of the computation power lies in the clock frequency of the GPU and the number of CUDA cores. The computation speed of the GPU is related to the clock frequency of the GPU and the number of threads the GPU can run concurrently is related to the number of CUDA cores. The graphics card here, GTX 1050, is an entry level GPU at the current moment. It has 640 CUDA cores and clock speed of 1455 MHz. In comparison, a relatively high level graphics card GeForce GTX 1080 has 2560 cores and a base clock speed of 1607 MHz. Hence, if a better GPU were used, it would be possible to further accelerate the full process.

As shown in the results, the larger the image size, the higher the difference in runtime in GPU and CPU. Although the GPU has a great advantage in computing large size Mueller matrix images, the size is limited by the graphic memory, which in this case is 2Gb. 2Gb of graphic memory is small for current standards; this may result in problems that involve large numbers of tensor stored in the memory for a series of computations. Especially in deep learning, training a network with dozens of layers requires moderate graphic memory to store a small batch of data. In personal experience, even a small batch of 64×64 Mueller matrix images can result in an out-of-memory problem on a 2Gb graphic card. However, in computing polarimetric parameters using polar decomposition on the GPU, the tensors required to be stored in graphic memory are four matrices ($\mathbf{M}$, $(\mathbf{m}_D)^{-1}$, $\mathbf{m'}(\mathbf{m'})^T$, $(\mathbf{m}_\Delta)^{-1}$) and 3 vectors ($\boldsymbol{D}$, $\boldsymbol{P}$, $\boldsymbol{P}_\Delta$). Hence, in the result, a Mueller matrix image up to 2000×2000 pixels can be computed without memory problems by using the 3×3 matrix notation.

The runtime of using CPU and original implementation to compute and decompose a 1000×1000 Mueller matrix is about 30s. GPU and accelerated implementation brought the runtime down to within 3s. This boost in performance significantly reduced waiting time either for analyzing a large database or analyzing for a set of Mueller matrix images in a clinical imaging device. It may be possible to further accelerate using better graphic cards and bring the processing time to a scale that is as small as the frame rate of the

imaging device. If the decomposition time can be reduced to shorter than the image acquisition time of the 16 images, it would be possible to achieve real time display of the polarimetric properties.

3.5 Conclusion

In this chapter, an accelerated implementation of polar decomposition was shown. This accelerated implementation decomposed the Mueller matrix in parallel with boosted performance in speed. Combining the power of GPU, one can achieve massive speed improvement over the original implementation. This acceleration has the potential to combine with a fast Mueller matrix imaging device to accomplish live display of the polarimetric parameters.

Chapter 4

Image segmentation of presumed amyloid deposits in Mueller matrix images of the retina

4.1 Introduction

For analyzing the properties of the presumed Aβ deposits in the retina from Chapter 2, it is important to separate deposits from surrounding retina in the Mueller matrix images. A binary mask needs to be created to extract the deposit region and exclude the surrounding retina when analyzing the deposit, and vice versa. Along with registration, the quality of segmentation can impact the accuracy of retardance and birefringence computation. Certainly, one can segment the image by hand using a good mouse and a great amount of patience. For a large database of deposits, the process needs to be automated. In addition, deposits do not have enough contrast in all the areas in just one polarimetric property. Even for linear retardance which has shown the largest contrast between the deposit and the surrounding retina, in some regions the contrast may be small. Combined analysis of multiple parameters or the Mueller matrix, the difference between the deposit and the retina can be enhanced. Hence, an automated segmentation method that uses the multidimensional polarimetric properties or Mueller matrix images is needed. In this chapter, several different segmentation methods including Otsu's method [92], K-means [93], Gaussian mixture model (GMM) [94], and Markov random field (MRF) [72] are used to segment the deposits from the background. In the result, their segmentation abilities are compared in different deposits with different polarimetric properties.

4.2 Methods

4.2.1 Segmentation algorithms

4.2.1.1 Otsu's method

Otsu's method [92] is an image segmentation method that applies to a single channel image. Suppose the values of the pixels in the image are normalized to m gray levels $I_j \in \{I_1, I_2, I_3, ...I_m\}$. For gray level I_j, the percentage of the total number of pixels in the image at this level is p_j. Assuming there is a gray level threshold I_t ($t \in \{1, 2, 3, ..., m\}$) that classifies the gray levels into two groups, this threshold determines the label $C_j \in \{1, 2\}$ of the all the pixels with gray level I_j in the image by

$$C_j = \begin{cases} 1, & I_j > I_t \\ 0, & I_j < I_t \end{cases} \tag{4.1}$$

To find the optimum threshold I_{t^*}, Otsu's method seeks to maximize the "between class variance" $\sigma_B^2(I)$ defined in Eq.4.2

$$\sigma_B^2(I_t) = \frac{(\mu_G P(I_t) - \mu(I_t))^2}{P(I_t)(1 - P(I_t))} \tag{4.2}$$

where μ_G is the average over all the pixels in the image. $P(I_t)$ and $\mu(I_t)$ are the percentage and weighted sum of pixels in the class having a value smaller than or equal to I_t. Thus, μ_G is written as

$$\mu_G = \sum_{j=1}^{m} p_j I_j \tag{4.3}$$

And $P(I_t)$ and $\mu(I_t)$ are given by

$$P(I_t) = \sum_{j=1}^{m} \{(\mathbb{1}\{I_j \leq I_t\}) * p_j\} \tag{4.4}$$

$$\mu(I_t) = \sum_{j=1}^{m} (\mathbb{1}\{I_j \leq I_t\}) * I_j p_j \tag{4.5}$$

where $\mathbb{1}\{\ \}$ is the indicator function which equals 1 when the statement inside the bracket is true. Otherwise it is zero. So the optimum threshold is then

$$I* = \arg\max_{I_t} \sigma_B^2(I_t) \tag{4.6}$$

Otsu's method automatically finds the best threshold to classify the pixels in a single channel image (for example a retardance image) into 2 classes. It can also be extended to classify pixels into more than 2 classes [95]. However, Otsu's method is defined on a single channel image. For a multichannel image, the methods of K-means clustering, GMM, MRF, introduced in the following sections, are used.

4.2.1.2 K-means clustering

Clustering algorithms help to group data into different classes based on their similarities. One clustering method used in this chapter is called K-means clustering [93]. Assume a group of data $\boldsymbol{x}_j = \{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_m\}$ are required to be separated into k classes. A k classes K-means clustering aims to find the optimal assignment of the label $C_j \in \{1, 2, 3..., k\}$ to each data point $\boldsymbol{x}_j$ so that the following objective function is minimized

$$J = \sum_{i=1}^{k} \sum_{j}^{m} \|\mathbb{1}\{C_j = i\} * (\boldsymbol{x}_j - \boldsymbol{\mu}_i)\|_2^2 \tag{4.7}$$

$\| \ \|_2$ is called $L2$ norm which represents the square root of the sum of the square of every element of the vector enclosed by this sign. $\boldsymbol{\mu}_i$ is the average of all the data points $\boldsymbol{x}_j$ whose label is $C_j = i$. It is computed as

$$\boldsymbol{\mu}_i = \frac{\sum_{j=1}^{m} \mathbb{1}\{C_j = i\}\boldsymbol{x}_j}{\sum_{j=1}^{m} \mathbb{1}\{C_j = i\}} \tag{4.8}$$

Hence, Eq.4.7 essentially aims to group the data points into clusters with centroid μ_i so that the data points in the same cluster are closer in distance to each other while data points in different clusters are further in distance from each other. To minimize Eq.4.7, an iterative approach listed here is used

1. Initialize k mean vectors $\{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3, ...\boldsymbol{\mu}_k\}$ by randomly selecting k data points from the data set. In our case, separating a deposit from its background means that $k = 2$.

2. Repeat until maximum iteration step

   (a) For every data point, compute $\|\boldsymbol{x}_j - \boldsymbol{\mu}_i\|_2^2$ with every $\boldsymbol{\mu}_i \in \{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3, ...\boldsymbol{\mu}_k\}$. Assign the label C_j to data $\boldsymbol{x}_j$ using

$$C_j = \arg\min_i \|\boldsymbol{x}_j - \boldsymbol{\mu}_i\|_2^2 \tag{4.9}$$

(b) Update the new mean vector $\boldsymbol{\mu}_i \in \{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3, ...\boldsymbol{\mu}_k\}$ using Eq.4.8

Once the maximum iteration step is reached or the mean vectors remain unchanged, compared to the last iteration, the final label for each data point is acquired. It has been shown in [96] that, when the dimension of the data point is 1, the objective function in Otsu's method is equivalent to the one in K-means clustering. Both methods depend on maximizing between class variance.

4.2.1.3 Gaussian Mixture Model

K-means clustering compares the absolute distance between a data point and the centroid of a class when assigning the label of this data point. Gaussian Mixture model (GMM) [94] modifies this "distance" as the probability in a weighted Gaussian distribution. This probabilistic approach gives a smoother decision boundary compared to K-means clustering [97]. A brief introduction of GMM is given in this subsection.
Suppose a k dimensional data set $\boldsymbol{x}_j = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3..., \boldsymbol{x}_m\}$ with m data points. The Gaussian Mixture Model (GMM) is based on fitting a mixture of Gaussian distributions onto the data. First, for a one component Gaussian distribution with mean and covariance of $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, the probability density $p(x|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is written as

$$p(x|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = (2\pi)^{\frac{k}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})) \tag{4.10}$$

where $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ are the mean and covariance of the ith Gaussian distribution. In GMM, there are n Gaussian components forming the distribution. Each Gaussian component is described by three parameters, the mean $\boldsymbol{\mu}_i$, the covariance $\boldsymbol{\Sigma}_i$ and a weighting of π_i in the distribution. Then, the probability of $\boldsymbol{x}_j$ can be written as the weighted sum of these n Gaussian components in the following equation

$$p(\boldsymbol{x}_j) = \sum_{i=1}^{n} \pi_i p(\boldsymbol{x}_j|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \tag{4.11}$$

Let $z_j = i \in \{1, 2, 3, ...n\}$ represent the ith Gaussian distribution which generates data point $\boldsymbol{x}_j$. Then the probability of this data point $\boldsymbol{x}_j$ generated by ith Gaussian distribution can be computed as the posterior probability using Bayes theorem in the following equation

$$p(z_j = i|\boldsymbol{x}_j) = \frac{p(z_j = i)p(\boldsymbol{x}_j|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)}{p_(\boldsymbol{x}_j)} \tag{4.12}$$

The prior probability $p(z_j = i)$ is the weighting π_i of the ith Gaussian component. So Eq.4.12 becomes

$$p(z_j = i|\boldsymbol{x}_j) = \frac{\pi_i p(\boldsymbol{x}_j|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)}{\sum_i^n \pi_i p(x_j|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)} \tag{4.13}$$

Then if the mean, covariance, and mixing coefficient $\{(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i, \pi_i)|1 \leq i \leq n\}$ of the GMM are known, the inferred label C_j of the data $\boldsymbol{x}_j$ is i in $\{1, 2, 3...n\}$ that has the maximum posterior probability as

$$C_j = \arg\max_i p(z_j = i|\boldsymbol{x}_j) \tag{4.14}$$

This determines the label of the data once the three parameters $\{(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i, \pi_i)|1 \leq i \leq n\}$ are known. However, in this case, these parameters are unknown. One needs to infer the parameters from the data using Maximum Likelihood (ML) estimation. The likelihood function of the model $\{(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i, \pi_i)|1 \leq i \leq n\}$ on the data set $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3..., \boldsymbol{x}_m\}$ is

$$\text{likelihood} = \prod_{j=1}^{m} p(\boldsymbol{x}_j) \tag{4.15}$$

Maximizing this likelihood function is equivalent to maximizing its logarithm $\ell(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \pi)$ as

$$\begin{aligned} \ell(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \pi) &= \log\Big(\prod_{j=1}^{m} p_{\boldsymbol{x}_j}\Big) \\ &= \sum_{j}^{m} \log\Big(\sum_{i}^{n} \pi_i p(\boldsymbol{x}_j|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)\Big) \end{aligned} \tag{4.16}$$

To maximize $\ell(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \pi)$, one can take the derivative with respect to $\{(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i, \pi_i)|1 \leq i \leq n\}$ and set it to 0. The results are

$$\boldsymbol{\mu}_i = \frac{\sum_{j=1}^{m} p(z_j = i|\boldsymbol{x}_j)\boldsymbol{x}_j}{\sum_{j=1}^{m} p(z_j = i|\boldsymbol{x}_j)} \tag{4.17}$$

$$\boldsymbol{\Sigma}_i = \frac{\sum_{j=1}^{m} p(z_j = i|\boldsymbol{x}_j)(\boldsymbol{x}_j - \mu_i)(\boldsymbol{x}_j - \mu_i)^T}{\sum_{j=1}^{m} p(z_j = i|\boldsymbol{x}_j)} \tag{4.18}$$

$$\pi_i = \frac{1}{m}\sum_{j=1}^{m} p(z_j = i|\boldsymbol{x}_j) \tag{4.19}$$

Once an initial guess of the labels are assigned to the data, which can be generated using K-means, the parameters of the model can be computed using Expectation Maximization method (EM) [98]. It is an iterative method which can be outlined as

1. Generate initial labels C'_j for each data $\boldsymbol{x}_j$ using K-means

2. Compute the initial guess of the parameters$\{(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i, \pi_i)|1 \leq i \leq n\}$ for the ith Gaussian component by computing the mean, covariance, and percentage of the data with the same initial label $C'_j = i$ in the whole data set

3. Repeat until maximum iteration step

   (a) Expectation step: compute the posterior probability using Eq.4.13

   (b) Maximization step: update model parameters $\{(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i, \pi_i)|1 \leq i \leq n\}$ using Eq.4.17,4.18,4.19

4. Infer the label C_j for data $\boldsymbol{x}_j$ using Eq.4.14

More specifically, in segmenting the deposit from the surrounding retina, the number of Gaussian components are 2 ($n = 2$). The number of data points m is the number of pixels in the image.

4.2.1.4 Markov Random Field (MRF)

MRF [72] is based on the assumption that the label of a pixel is influenced by two factors. First, its likelihood in a distribution created by all the pixels in the image with the same label. Second, if one label has a bigger population in the pixels surrounding this pixel of interest, this label will have a larger probability multiplied onto the probability created by the first factor. Representing the observed image as Y (each pixel having k channels) and the label image as X (n different labels in total), the posterior probability of the label image given the observed image can be written using Bayes theorem as

$$\mathrm{P}(X|Y) \propto \mathrm{P}(Y|X)\mathrm{P}(X) \tag{4.20}$$

For a pixel $\boldsymbol{y}$ (a k-dimensional vector), its probability of assignment to a label $x \in \{1, 2, 3...n\}$ is written as

$$\mathrm{P}(X = x|Y = \boldsymbol{y}) \propto \mathrm{P}(Y = \boldsymbol{y}|X = x)\mathrm{P}(X = x) \tag{4.21}$$

where $\mathrm{P}(Y = \boldsymbol{y}|X = x)$ is the likelihood of this pixel $\boldsymbol{y}$ assigned to label x while $\mathrm{P}(X = x)$ is the prior probability of the labels. First, the likelihood measures the probability of $\boldsymbol{y}$ in the distribution formed by all the pixels with the same label x in the image. The most widely used distribution representing this likelihood is the Gaussian distribution [99]. Thus, the likelihood is written as a multidimensional Gaussian distribution whose form is

$$\mathrm{P}(Y = \boldsymbol{y}|X = x) = (2\pi)^{\frac{-k}{2}}|\boldsymbol{\Sigma}_x|^{-\frac{1}{2}}\exp(-\frac{1}{2}(\boldsymbol{y} - \boldsymbol{\mu}_x)^T\boldsymbol{\Sigma}_x^{-1}(\boldsymbol{y} - \boldsymbol{\mu}_x)) \tag{4.22}$$

where $\boldsymbol{\mu}_x$ and $\boldsymbol{\Sigma}_x$ are the mean and covariance of the pixels with label x in the image. The prior distribution $\mathrm{P}(X = x)$ is modeled as a conditional distribution with Markov property, which also gives the name to this method. The Markov property dictates that the label of a pixel is dependent on the label of its neighboring pixels connecting to it. Representing the label of these neighboring pixels as N_x, the prior probability $\mathrm{P}(X = x)$ becomes

$$\mathrm{P}(X = x) = \mathrm{P}(X = x|N_x) \tag{4.23}$$

In MRF, this probability is modeled as the probability in a Gibbs distribution. It has been proven in [100] that label maps with the Markov property follow the Gibbs distribution. If the labels of the neighboring pixels N_x have a metric $U(x)$ that quantifies its homogeneity, the Gibbs probability is defined as

$$\mathrm{P}(X = x) = \frac{1}{Z}\exp(-\beta U(x)) \tag{4.24}$$

where Z is the normalizing factor to normalize the probability and β is a constant which weights the influence of the labels of the neighboring pixels. The metric $U(x)$ for N_x used here is given by the Potts model [101] which can be written as

$$U(x) = \sum_{x' \epsilon N_x} (1 - \delta(x, x')) \tag{4.25}$$

where $\delta(x, x')$ is the Kronecker delta function so that $(1 - \delta(x, x'))$ is

$$1 - \delta(x, x') = \begin{cases} 0, & x = x' \\ 1, & x \neq x' \end{cases} \tag{4.26}$$

Hence, rewriting Eq.4.21 using Eq.4.22 and Eq.4.24 gives

$$\begin{aligned} \mathrm{P}(X = x|Y = \boldsymbol{y}) \propto & (2\pi)^{\frac{-k}{2}} |\boldsymbol{\Sigma}_x|^{-\frac{1}{2}} \exp(-\frac{1}{2}(\boldsymbol{y} - \boldsymbol{\mu}_x)^T \boldsymbol{\Sigma}_x^{-1}(\boldsymbol{y} - \boldsymbol{\mu}_x)) \\ & \exp(-\beta U(x))/Z \end{aligned} \tag{4.27}$$

To the simplify this expression, one can take the logarithm of the above equation which results in

$$\begin{aligned} \log(\mathrm{P}(X = x|Y = \boldsymbol{y})) \propto & -(\frac{k}{2}\log(2\pi) + \log(Z)) + [-\frac{1}{2}(\boldsymbol{y} - \boldsymbol{\mu}_x)^T \boldsymbol{\Sigma}_x^{-1}(\boldsymbol{y} - \boldsymbol{\mu}_x) \\ & - \beta \sum_{x' \epsilon N_x} (1 - \delta(x, x'))] \end{aligned} \tag{4.28}$$

where the term $(\frac{k}{2}\log(2\pi)+\log(Z))$ is constant for all labels. Thus the final log probability is written as

$$\log(\mathrm{P}(X=x|Y=\boldsymbol{y})) \propto -\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}_x)^T\boldsymbol{\Sigma}_x^{-1}(\boldsymbol{y}-\boldsymbol{\mu}_x) - \beta\sum_{x'\in N_x}(1-\delta(x,x')) \tag{4.29}$$

Then the segmentation problem is equivalent to finding the label $x^* \in \{1,2,3...n\}$ with the maximum posterior probability

$$x^* = \arg\min_x \ -\frac{1}{2}(\boldsymbol{y}-\boldsymbol{\mu}_x)^T\boldsymbol{\Sigma}_x^{-1}(\boldsymbol{y}-\boldsymbol{\mu}_x) - \beta\sum_{x'\in N_x}(1-\delta(x,x')) \tag{4.30}$$

This is done by assigning the label x to the pixel y, where x takes on the values in $\{1,2,3...n\}$ and using Eq.4.30 to compute their posterior probability. Then pick the label with the maximum posterior probability x^* as the label of this pixel. In particular, the method in which the MRF is implemented in this Chapter is called the Iterated Conditional Modes (ICM) [72]. The procedure is outlined as

1. Generate an n level initial mask using K-means clustering
2. Iterate until maximum iteration step
   (a) For every pixel, while fixing the labels of all other pixels, compute the log probability of assigning the label of this pixel to all labels from 1 to n using Eq.4.29.
   (b) Using Eq.4.30, assign the label with the minimum log probability to each pixel.

4.2.2 Image preparation

The images used in testing segmentation methods were acquired from the Mueller matrix microscope used in Chapter 2. Images of 771 locations with amyloid deposits from 22 subjects all with high likelihood of AD were used in this chapter to test the accuracy of the methods. For testing Otsu's method, single channel images (linear retardance R_L, depolarization power Δ and linear anisotropy A_L) were used. Their derivations are introduced in Chapter 1, 3. So, Otsu's method was evaluated on each of the 3 images separately. For testing K-means, GMM and MRF, multichannel images were used. To form a multichannel image, 15 elements from the normalized Mueller matrix were used (not including the

M_{00} intensity image). These 15 elements were first reshaped into a vector I_M and then standardized using Eq.4.31

$$I'_M = \frac{I_M - \mu(I_M)}{\sigma(I_M)} \tag{4.31}$$

where $\mu(I_M)$ and $\sigma(I_M)$ are the mean and standard deviation of each channel. The K-means, GMM and MRF were then evaluated on this standardized 15 channel image I'_M. The resulting 6 different approaches to segment the deposits from the surrounding retina are listed in Table.4.1

Algorithm	Image analized
Otsu	R_L
Otsu	Δ
Otsu	A_L
K-means	Normalized Mueller matrix (Standardized by Eq.4.31)
GMM	Normalized Mueller matrix (Standardized by Eq.4.31)
MRF	Normalized Mueller matrix (Standardized by Eq.4.31)

Table 4.1: Table of segmentation methods used in this Chapter listed with algorithm and image type analized.

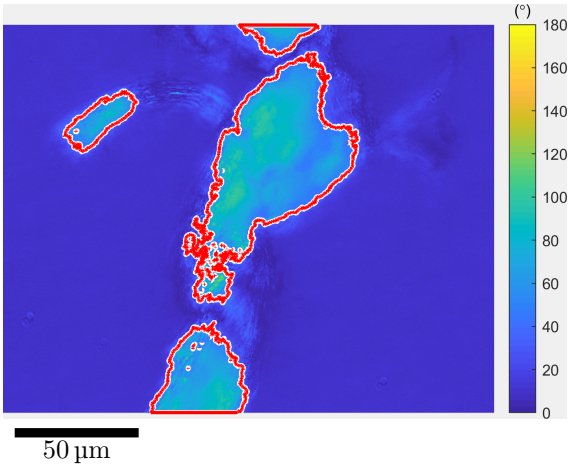
The ground truth of the mask of each deposit was created by choosing the best segmentation result and further assisted by fine tuning the parameters of the segmentation method when necessary. The resulting mask of each of the methods listed in Table.4.1 was evaluated by comparing to the ground truth.

4.3 Results

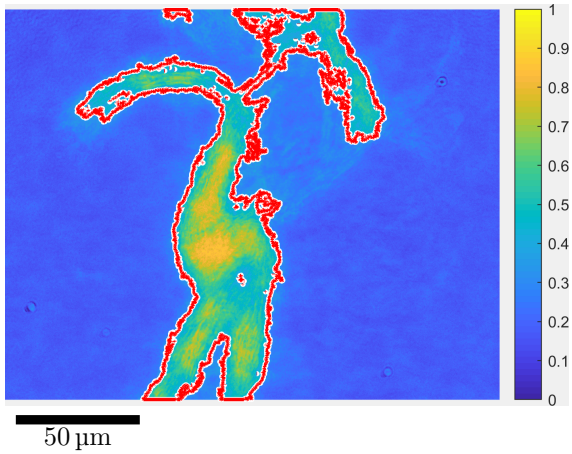
4.3.1 Results of Otsu's method and K-means clustering

The segmentation results of using Otsu's method on R_L, Δ, and A_L images, represented by Otsu(R_L), Otsu(Δ, Otsu(A_L)) respectively, and the result of using K-means clustering on the multichannel images are shown in Fig.4.1. As shown in Fig.4.1, the segmentation results on R_L image and A_L image are similar. Otsu's method on these two images cut off at the region with high R_L and A_L values while some regions, which have lower R_L and A_L values which appear by eye to be higher than the surround and part of the deposit, are ignored. Otsu's method on the Δ image incorporates these low R_L and A_L regions as part

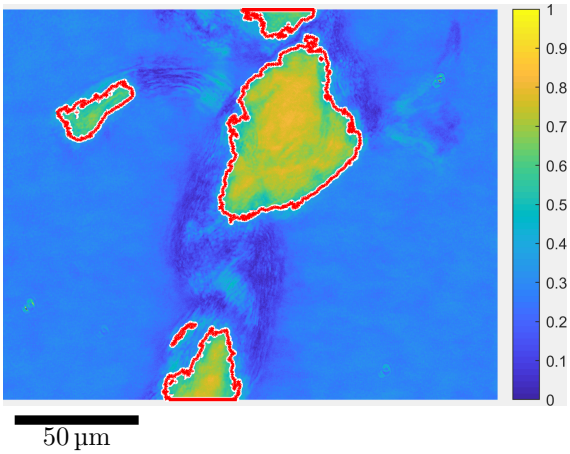
of the deposit. However, some regions which appear to be part of the upper portion of the deposits are left out in the segmentation of the Δ image. As a result, Otsu's method gives differing, apparently inaccurate results or each of the 3 images. The result of K-means clustering shown in the 4th image in Fig.4.1 looks like the combination of the result from Otsu's method on the R_L, A_L and Δ images. It is the most accurate result among the four approaches on this deposit. However, it is not an ideal segmentation result since it ignores some regions towards the top of the image which appear to be part of the deposit.



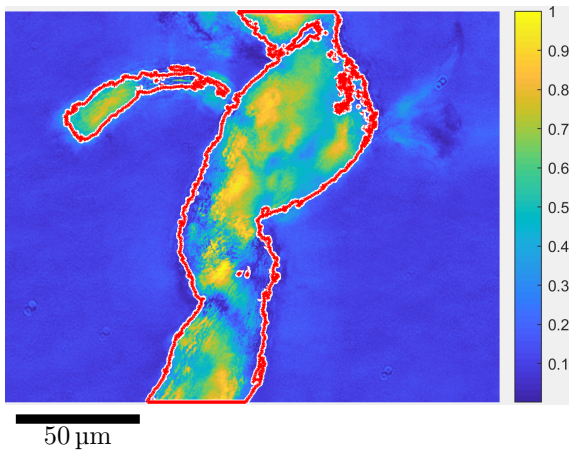
(a) Otsu(R_L)



(b) Otsu(Δ)



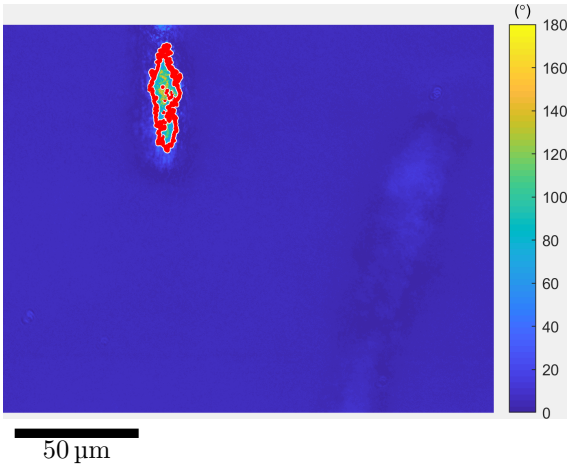
(c) Otsu(A_L)



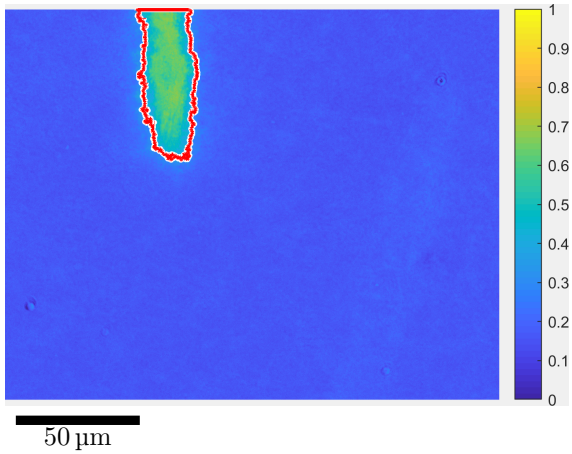
(d) K-means

Figure 4.1: Segmentation by Otsu's method on: (a) R_L image, (b) Δ image, and (c) A_L image of one deposit. (d) Segmentation by K-means clustering on the same deposit displayed on a contrasted R_L image. Red contours in all images represent the segmented mask.

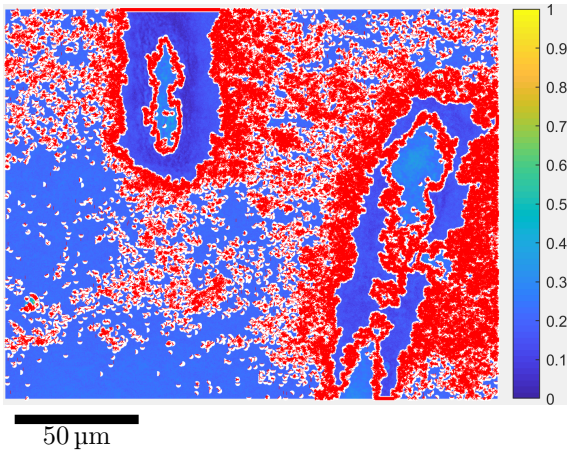
Another example is shown in Fig.4.2. In Fig.4.2, all methods show poor performance in segmenting the two deposits. Otsu(R_L), Otsu(Δ) and K-means clustering only segment the one deposit at the top of the image. Although Otsu(A_L) recognizes that there are two deposits in the image, it also includes a large part of the retina as deposit but excludes the regions of the deposit with low A_L.



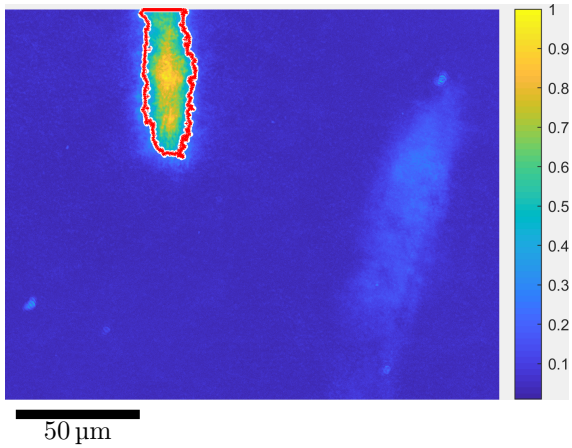
(a) Otsu(R_L)



(b) Otsu(Δ)



(c) Otsu(A_L)

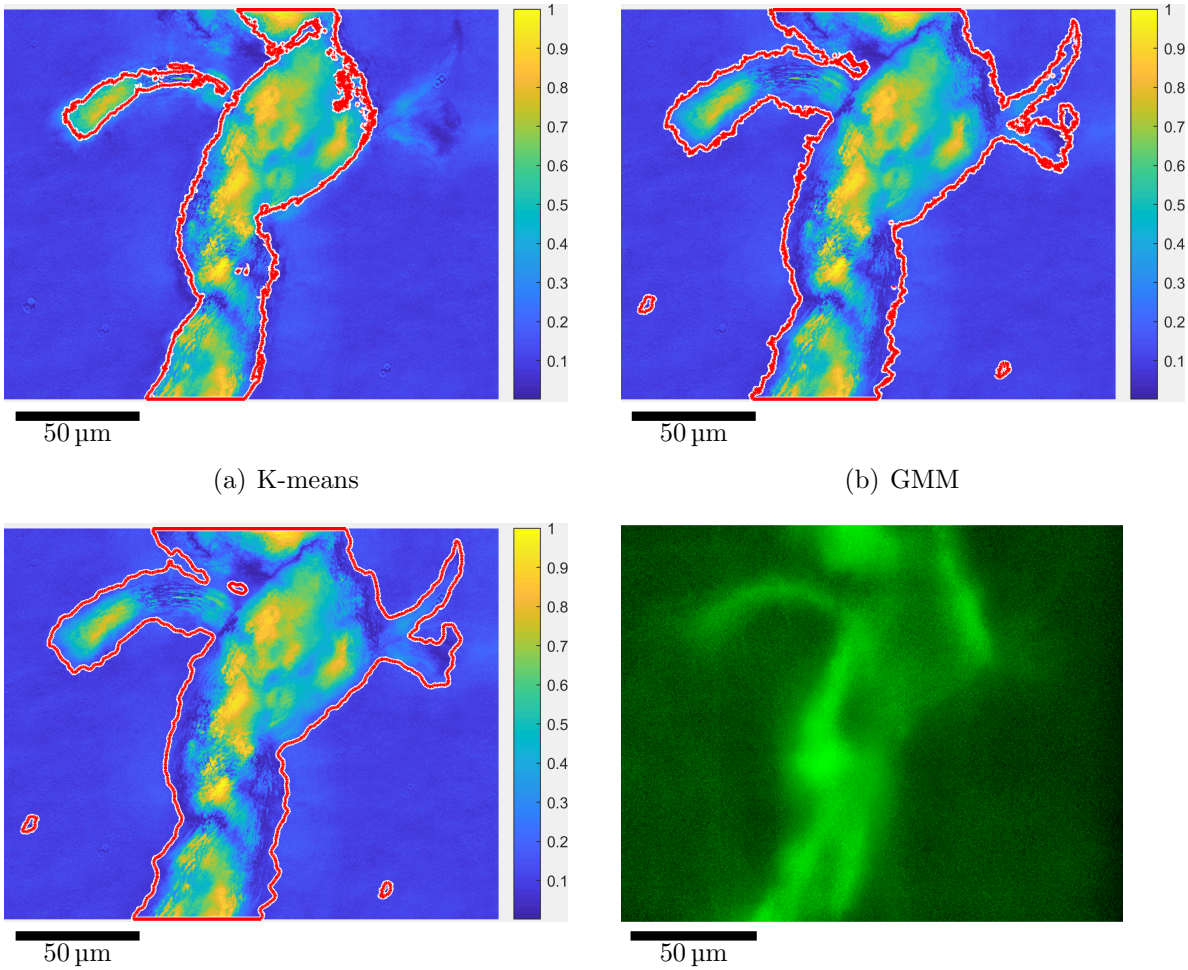


(d) K-means

Figure 4.2: Segmentation by Otsu's method on: (a) R_L image, (b) Δ image, and (c) A_L image of two deposits in one image. (d) Segmentation by K-means clustering on the same deposits displayed on a contrasted R_L image. Red contours in all images represent the segmented mask.

4.3.2 Results of GMM and MRF

The deposit in Fig.4.1 with the segmentation by K-means clustering compared to that by GMM and MRF is shown in Fig.4.3.



(a) K-means

(b) GMM

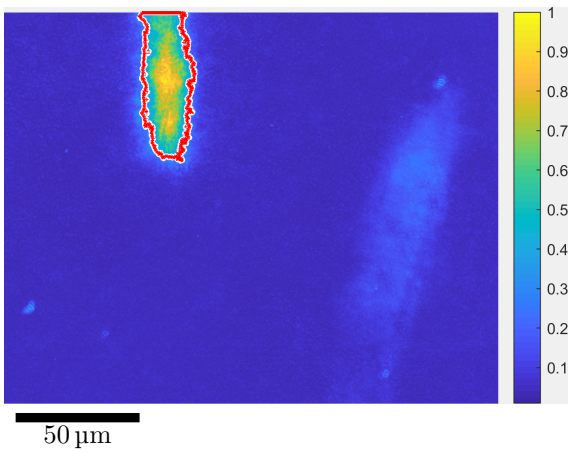
(c) MRF

(d) Fluorescence

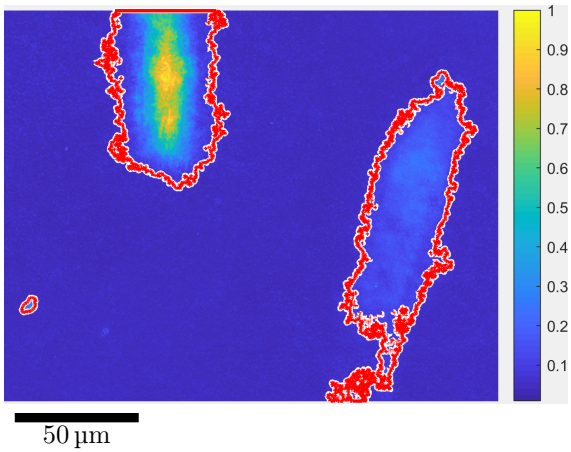
Figure 4.3: Segmentation quality of (a) K-means clustering, (b) GMM, (c) MRF on the same deposit as Fig.4.2. Red contours in all images represent the segmented mask. The segmentation results are displayed on a contrasted R_L image. (d) is the fluorescence image.

Compared to the K-means result shown in Fig.4.3.(a), both the GMM and MRF successfully segment the whole deposit which is visible in the contrasted R_L image. The upper right part of the deposit, as shown in the contrasted R_L image (background of Fig.4.3.(a),(b),(c)), appears in this image as well to be part of the deposit which might be better to be included in the mask. The K-means clustering result in Fig.4.3.(a) fails to segment this part while both masks in Fig.4.3.(b) and (c) successfully classify this region

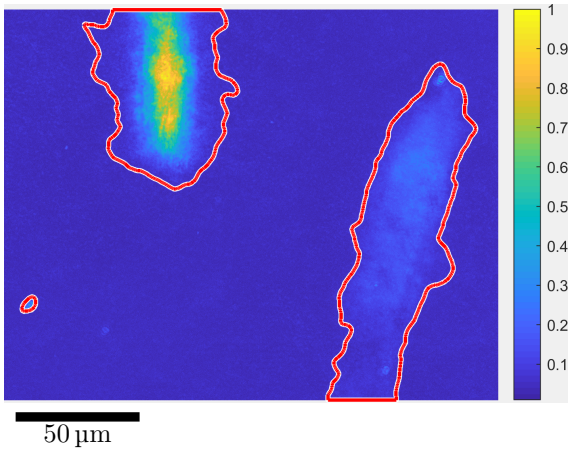
as part of the deposit. The segmented contours in both masks closely follow the boundary of the deposit which can be considered as an excellent segmentation result.
The results of using GMM and MRF to segment the image in Fig.4.2 are also shown in Fig.4.4



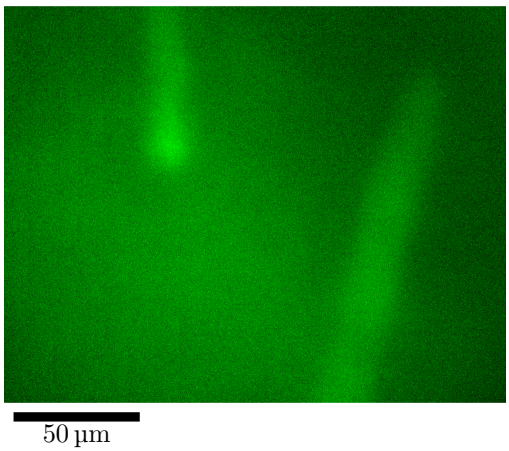
(a) K-means



(b) GMM



(c) MRF



(d) Fluorescence

Figure 4.4: Segmentation by (a) K-means clustering, (b) GMM, (c) MRF on the same deposit as Fig.4.2. Red contours in all images represent the segmented mask. The segmentation results are displayed on a contrasted R_L image. (d) is the fluorescence image.

Compared to K-means clustering, the results of GMM and MRF shown in Fig.4.4.(b),

(c) indicate that the two methods are capable of detecting the deposit visible to the eye in the bottom of the image. In Fig.4.4.(d), the fluorescence image further confirmed the presence of the second deposit. The GMM in Fig.4.4.(b) segments a "thinner" mask compared to the MRF result. When viewing the fluorescence image in Fig.4.4.(d), the MRF mask seems to incorporates the full shape of the deposit. Hence, the MRF mask appears by eye to be more accurate in defining the boundary of the deposit than the GMM mask.

4.3.3 Overall performance

The overall performance of the 6 methods are shown in Table.4.2. The success criterion for segmentation is defined as over 90% of the pixels in the image are correctly classified compared to the ground truth. The performance of MRF after 30 iterations achieved the best success rate of 96.8% among all methods. A slightly lower success rate of 95.1% was achieved by GMM but it can finish within 10s per image compared to 45 s for the MRF. Although having the fastest speed in segmentation, both K-means and Otsu's methods performed poorly with a success rate lower than 80 %.

Rank	Method	Success rate	Runtime (per image)
1	MRF (30 iterations)	96.8%	45 s
2	GMM	95.1%	9 s
3	K-means	73.8%	2 s
4	Otsu(R_L)	59.7%	0.5 s
5	Otsu(A_L)	51.8%	0.5 s
6	Otsu(Δ)	43.7%	0.5 s

Table 4.2: Table of performance of the 6 methods ranked by success rate.

4.4 Discussion

The result shown in section 4.3 is an evaluation of all the methods listed in Table.4.1. Both GMM and MRF achieve a success rate of more than 95%. Hence, both methods can be used for automatic segmentation of a large data base to replace hand crop segmentation. Note that manually checking the segmentation quality is still required. However, compared to segmenting 771 deposits by hand, this method is less laborious, while simultaneously

more time and cost effective. In addition, the structure of the deposit does not show in just one polarimetric image. Deposits may have regions with low R_L which can have high Δ properties. Hence, manual segmentation requires alternating between several polarimetric images for a complete understanding of the structure of the deposit. Also, if more than one person is segmenting a data set, the uncertainty of the segmented mask in the data set can increase, since manual segmentation is inherently subjective.

Among the 3 images of polarization properties used to test Otsu's method, the linear retardance is the most consistent and strongest feature of an amyloid deposit. Hence, the segmentation accuracy using Otsu's method achieves the best result when using the linear retardance image R_L as the feature image. The depolarization image Δ does not have good contrast between the deposit and the background thus more than half of the segmented masks using Otsu's method on these images have poor quality. The K-means clustering, which uses a multichannel image, achieved much better performance than Otsu's method on the single channel image. Since Otsu's methods and K-means clustering are both based on maximizing the between class variance [96], the improved performance by K means clustering indicates that segmenting amyloid deposits using a multichannel image is better than using single channel image of one polarimetric parameter. However, the K-means clustering only reaches 70% of success rate which is much lower than GMM which also uses a multichannel image.

When both methods use the same multichannel image, the high accuracy of the GMM method vs K-means clustering indicates that the Mueller matrices of the deposit and background are more likely to be generated by two distinct Gaussian distributions with different means and covariances. The MRF and GMM both showed great performance in segmenting the deposits while GMM is much faster than the MRF. The mask produced by MRF is generally wider than the GMM. Since the MRF uses the information of the neighboring pixels, the boundary pixels are likely to be incorporated into the mask if their polarization features are closer to those of the deposit than the retina. Although MRF achieves the highest accuracy, a slower speed is the main drawback of this method. However, this could potentially be improved by adopting a Hidden Markov Chains Model (HMCM) or a Hidden Hierarchical Markov Model (HHMM) which are both non iterative ways of segmenting the image with faster speed [74] .

All the segmentation methods presented in this chapter are unsupervised methods which do not depend on any training examples to train the classifier. To further improve the segmentation result, a supervised approach using training examples to train a classifier could be applied. This classifier should consider the spatial features of the deposit as well as the polarization properties to separate the deposit and the retina. One state of the art segmentation method that uses both the spatial and channel features of the image is semantic segmentation using a Convolutional Neural Network (CNN). Most CNN

architectures require a large number of training examples to successfully train the CNN to make ideal predictions. This is not applicable in this case since a large number of segmented data is not available. However, one type of CNN architecture called U-net [102] can achieve great accuracy using only a small number of training examples. Thus, this architecture is a good method to try on the Mueller matrix images to improve the segmentation accuracy.

4.5 Conclusion

In this Chapter, Otsu's method, K-means clustering, GMM and MRF were evaluated in terms of segmentation quality and efficiency in segmenting amyloid deposits from the background retina using the Mueller matrix images. MRF showed the best overall performance in segmenting the deposits accurately, but took a relatively long time to segment images compared to the GMM method which was almost as accurate. The results showed the merits of adopting an automated segmentation method which can significantly reduce labour while achieving good accuracy. Further improvements may perhaps be made by a supervised approach such as U-net based on CNN.

Chapter 5

Conclusion and future directions

5.1 Summary of Results

In this thesis, three different but interconnected research projects are presented. First, in Chapter 2, the birefringence of amyloid deposits in the retinas of subjects with AD were compared to pure Aβ deposits grown on glass slides. To successfully analyze the properties of the deposits, the polarization properties were measured using MMP and were calculated using polar decomposition. To optimize and speed up this process, in Chapter 3, a fast implementation of the polar decomposition as well as fast computation of the Mueller matrix were achieved using a GPU. To compute the birefringence and analyze the polarimetric properties of the deposit, image segmentation for separating the deposit from the surrounding retina or buffer was conducted. So in Chapter 4, several different methods of automatically segmenting the images of amyloid were used and analyzed. The image segmentation used in Chapter 2 was the MRF method which achieved the highest accuracy in Chapter 4.

In Chapter 2, the birefringence of the amyloid deposits, both retinal and pure with similar thicknesses, were imaged and computed using the combination of MMP, CLSM and image processing techniques (image registration and image segmentation). In both types of deposits, the linear retardance values increased with thickness in the thinner regions then plateaued and decreased at higher thickness, resulting in lower birefringence at higher thickness. The slope of the fit of linear retardance as a function of thickness in the linear region in retinal and pure deposits showed no significant difference. This result suggests structural similarity between the retinal and pure deposits which indicates that the retinal deposits are composed predominantly of Aβ fibrils. The result of low birefringence in

high thickness regions also suggests only short range order of the Aβ fibrils in the deposit. The measured average linear birefringence and retardance values are much higher than the reported birefringence of RNFL [50]. This means that retinal deposits are expected to be visible against the RNFL of the retina. Additionally, the measured birefringence values of of retinal deposits are also higher than the senile plaques measured by PSOCT [29] even though senile plaques are much thicker than retinal deposits. The thicker pure deposits show birefringence values similar to previously measured in senile plaques, likely because of long range disorder. The high birefringence value of the retinal deposits will be useful in future clinical detection using a patented label free method [88].

To accelerate the speed of computing, the Mueller matrix and polar decomposition, a GPU was used and the algorithm was implemented in an optimized way in Chapter 3. When both were computing on the same CPU, the optimized implementation was more than 5 times faster than the original implementation. The GPU then further accelerated the optimized implementation to 5 times faster than computing it on the CPU. The decomposition of an image of size 1000×1000 was shortened from more than 30 s to within 2 s. The improvement in speed in computing and decomposing the Mueller matrix assisted by GPU would greatly reduce both the waiting time in clinical polarization imaging and processing time of a large database. The result also suggested the possibility of integrating better GPU with the accelerated polar decomposition into a clinical device to achieve fast real time display of the polarization properties.

Several segmentation methods including Otsu's method, K-means clustering, GMM, and MRF were examined in Chapter 4. Both Otsu's method and K-means clustering had trouble in segmenting out the full morphology of the deposit. Otsu's method performed poorer than K-means clustering while K-means clustering didn't reach over an 80% success rate. GMM and MRF both achieved over a 95% of success rate. The MRF showed the highest success rate in segmenting the deposit from the background retina, however the runtime of 30 times of iterations is around 45 s. GMM also achieved a high success rate of over 95% while performing much faster than the MRF. Since both GMM and MRF showed good segmentation result, these two methods are suitable for automatic segmentation of a large database to replace hand cropping segmentation.

In conclusion, the measured birefringence properties brought insight into the structure of the retinal amyloid deposits. With the help of GPU, the morphology and polarization properties of the deposits under polarized light can be acquired at a much faster speed. Image segmentation for analyzing the polarimetric properties of the deposits and retina can be conducted automatically in large quantities with high accuracy using GMM or MRF.

5.2 Future directions

The work presented in this thesis can be extended in several ways.

1. Birefringence of different types of proteins:
   The method of measuring birefringence of amyloid deposits from the retina of AD subjects shown in Chapter 2 can be extended to measuring amyloid and other protein deposits from other diseases. For example, amyloid is also found in the retina of subjects with cerebral Malaria [103] and other brain trauma, in a different morphology than in AD. Measuring linear birefringence of these deposits could potentially help to increase the understanding of their structure.

2. Mueller matrix decomposition on the GPU:
   The Mueller matrix decomposition shown in Chapter 3 accelerated on a GPU is the polar decomposition method. Other decompositions such as symmetric decomposition and differential decomposition can also be accelerated on the GPU for comparing their speed to the polar decomposition. For computing on the GPU, the original algorithm is required to be organized in the parallel. In addition, the GPU computation used in Chapter 3 is conducted using the MATLAB Parallel Computing toolbox [91]. Currently, many machine learning libraries such as Tensorflow [104] and Pytorch [105] come with good GPU computation functions. These libraries can be used to conduct the decomposition and compare with the MATLAB parallel computing toolbox.

3. A path towards machine learning:

   (a) Improving segmentation results using supervised learning:
   The segmentation used in Chapter 4 is an unsupervised approach. It reached its limit when encountering deposits with low polarization values. To successfully segment these deposits, a supervised approach could be tested to train a classifier to have a good ability in recognizing the channel information as well as the spatial information of these deposits.

   (b) Diagnostic model for AD with high accuracy:
   Using a database to create a machine learning framework for diagnosing AD by combining polarimetric properties of deposits, the background retina, number of deposits, size of the deposits, and other available parameters. Support Vector Machine (SVM) and simple multilayer feedforward neural networks are two off the shelf methods to implement at first.

References

[1] Warren W Barker, Cheryl A Luis, Alice Kashuba, Mercy Luis, Dylan G Harwood, David Loewenstein, Carol Waters, Pat Jimison, Eugene Shepherd, Steven Sevush, Neil Graff-Radford, Douglas Newland, Murray Todd, Bayard Miller, Michael Gold, Kenneth Heilman, Leilani Doty, Ira Goodman, Bruce Robinson, Gary Pearl, Dennis Dickson, and Ranjan Duara. Relative frequencies of Alzheimer disease, Lewy body, vascular and frontotemporal dementia, and hippocampal sclerosis in the State of Florida Brain Bank. *Alzheimer disease and associated disorders*, 16(4):203–12, 2002.

[2] Alzheimer's Association. 2018 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, 14(3):367–429, mar 2018.

[3] Martin Prince, Renata Bryce, Emiliano Albanese, Anders Wimo, Wagner Ribeiro, and Cleusa P. Ferri. The global prevalence of dementia: A systematic review and metaanalysis. *Alzheimer's & Dementia*, 9(1):63–75.e2, jan 2013.

[4] World Health Organization. and Alzheimer's Disease International. *Dementia : A Public Health Priority.* World Health Organization, 2012.

[5] Philip Scheltens, Kaj Blennow, Monique M B Breteler, Bart de Strooper, Giovanni B Frisoni, Stephen Salloway, and Wiesje Maria Van der Flier. Alzheimer's disease. *The Lancet*, 388(10043):505–517, 2016.

[6] E Genin, D Hannequin, D Wallon, K Sleegers, M Hiltunen, O Combarros, M J Bullido, S Engelborghs, P De Deyn, C Berr, F Pasquier, B Dubois, G Tognoni, N Fiévet, N Brouwers, K Bettens, B Arosio, E Coto, M Del Zompo, I Mateo, J Epelbaum, A Frank-Garcia, S Helisalmi, E Porcellini, A Pilotto, P Forti, R Ferri, E Scarpini, G Siciliano, V Solfrizzi, S Sorbi, G Spalletta, F Valdivieso, S Vepsäläinen, V Alvarez, P Bosco, M Mancuso, F Panza, B Nacmias, P Bossù, O Hanon, P Piccardi, G Annoni, D Seripa, D Galimberti, F Licastro, H Soininen, J-F Dartigues, M I

Kamboh, C Van Broeckhoven, J C Lambert, P Amouyel, and D Campion. APOE and Alzheimer disease: a major gene with semi-dominant inheritance. *Molecular Psychiatry*, 16(9):903–907, sep 2011.

[7] Eric Karran, Marc Mercken, and Bart De Strooper. The amyloid cascade hypothesis for Alzheimer's disease: an appraisal for the development of therapeutics. *Nature Reviews Drug Discovery*, 10(9):698–712, sep 2011.

[8] Thomas L Williams, Benjamin R G Johnson, Brigita Urbanc, A Toby A Jenkins, Simon D A Connell, and Louise C Serpell. Aβ42 oligomers, but not fibrils, simultaneously bind to and cause damage to ganglioside-containing lipid membranes. *The Biochemical journal*, 439(1):67–77, oct 2011.

[9] Scott A. Small and Karen Duff. Linking Aβ and Tau in Late-Onset Alzheimer's Disease: A Dual Pathway Hypothesis. *Neuron*, 60(4):534–542, 2008.

[10] William E. Klunk, Henry Engler, Agneta Nordberg, Yanming Wang, Gunnar Blomqvist, Daniel P. Holt, Mats Bergström, Irina Savitcheva, Guo Feng Huang, Sergio Estrada, Birgitta Ausén, Manik L. Debnath, Julien Barletta, Julie C. Price, Johan Sandell, Brian J. Lopresti, Anders Wall, Pernilla Koivisto, Gunnar Antoni, Chester A. Mathis, and Bengt Långström. Imaging Brain Amyloid in Alzheimer's Disease with Pittsburgh Compound-B. *Annals of Neurology*, 55(3):306–319, 2004.

[11] A. Anoop, Pradeep K. Singh, Reeba S. Jacob, and Samir K. Maji. CSF Biomarkers for Alzheimer's Disease Diagnosis. *International Journal of Alzheimer's Disease*, 2010(Table 1):1–12, 2010.

[12] C. Ising, M. Stanley, and D. M. Holtzman. Current thinking on the mechanistic basis of Alzheimer's and implications for drug development. *Clinical Pharmacology and Therapeutics*, 98(5):469–471, 2015.

[13] a Cronin-Golomb, S Corkin, J F Rizzo, J Cohen, J H Growdon, and K S Banks. Visual dysfunction in Alzheimer's disease: relation to normal aging. *Annals of neurology*, 29:41–52, 1991.

[14] Fatmire Berisha, Gilbert T. Feke, Clement L. Trempe, J. Wallace McMeel, and Charles L. Schepens. Retinal abnormalities in early Alzheimer's disease. *Investigative Ophthalmology and Visual Science*, 48(5):2285–2289, 2007.

[15] Carol Yim, Lui Cheung, Yi Ting Ong, Saima Hilal, M Kamran Ikram, Sally Low, Yi Lin Ong, N Venketasubramanian, Philip Yap, Dennis Seow, Christopher Li, Hsian

Chen, and Tien Yin Wong. Retinal Ganglion Cell Analysis Using High-Definition Optical Coherence Tomography in Patients with Mild Cognitive Impairment and Alzheimer's Disease. *Journal of Alzheimer's Disease Retinal Ganglion Cell Analysis in MCI & AD*, 45(1):45–56, 2015.

[16] Carol Yim Lui Cheung, Yi Ting Ong, M. Kamran Ikram, Shin Yeu Ong, Xiang Li, Saima Hilal, Joseree Ann S Catindig, Narayanaswamy Venketasubramanian, Philip Yap, Dennis Seow, Christopher P. Chen, and Tien Yin Wong. Microvascular network alterations in the retina of patients with Alzheimer's disease. *Alzheimer's and Dementia*, 10(2):135–142, 2014.

[17] MCW Campbell, L Gowing, Y Choi, and Z Leonenko. Imaging of amyloid-beta deposits in the post-mortem retina in alzheimer's disease. In *Investigative Ophthalmology & Visual Science*, page 50, Ft. Lauderdale, 2010. ARVO.

[18] Nadav J. Hart, Yosef Koronyo, Keith L. Black, and Maya Koronyo-Hamaoui. Ocular indicators of Alzheimer's: exploring disease in the retina. *Acta Neuropathologica*, 132(6):767–787, 2016.

[19] Cheng-Ying Ho, Juan C. Troncoso, David Knox, Walter Stark, and Charles G. Eberhart. Beta-Amyloid, Phospho-Tau and Alpha-Synuclein Deposits Similar to Those in the Brain Are Not Identified in the Eyes of Alzheimer's and Parkinson's Disease Patients. *Brain Pathology*, 24(1):25–32, jan 2014.

[20] AN Begum, A Kvanta, S Seregard, MK Staples, MJ Radeke, DH Anderson, Shawn J. Kile, Austin Blanco, Dieu-Trang Fuchs, Adeel Ashfaq, Sally Frautschy, Gregory M. Cole, Carol A. Miller, David R. Hinton, Steven R. Verdooner, Keith L. Black, and Maya Koronyo-Hamaoui. Retinal amyloid pathology and proof-of-concept imaging trial in Alzheimer's disease. *J Pharmacol Exp Ther*, 326(1):196–208, 2017.

[21] P Divry. Etude histochimique des plaques séniles. *J. Belge. Neurol. Psych.*, 27:643, 1927.

[22] H. Bennhold. Eine specifische amyloidfärbung mit kongorot. *Münchener Medizinische Wochenschrift*, 69:1537, 1922.

[23] Ritu Khurana, Vladimir N. Uversky, Liza Nielsen, and Anthony L. Fink. Is Congo Red an Amyloid-specific Dye? *Journal of Biological Chemistry*, 276(25):22715–22721, 2001.

[24] Lee-Way Jin, Kacey A Claborn, Miki Kurimoto, Morten A Geday, Izumi Maezawa, Faranak Sohraby, Marcus Estrada, Werner Kaminksy, and Bart Kahr. Imaging linear birefringence and dichroism in cerebral amyloid pathologies. *Proceedings of the National Academy of Sciences of the United States of America*, 100(26):15294–8, dec 2003.

[25] MCW Campbell. Method and system for imaging amyloid beta in the retina of the eye in association with alzheimer's disease, may 2011.

[26] Michael Tokiyoshi Hamel, Laura Emptage, David DeVries, Catalina Oliveros, Theodore Chow, Namrata Shah, Chris Cookson, Marsha L Kisilak, and Melanie C W Campbell. Polarization properties of amyloid deposits in the retinas of an animal model of Alzheimer's disease differ in those with and without cognitive impairment. In *Investigative Ophthalmology & Visual Science*, volume 57, pages 2216–2216. C.V. Mosby Co, sep 2016.

[27] Tao Jin, Laura Emptage, David DeVries, and Melanie C. Campbell. Mapping the birefringence of amyloid deposits found in retinas in association with Alzheimer's disease. *Frontiers in Optics 2016 (2016), paper FTh5D.4*, page FTh5D.4, 2016.

[28] Melanie C. Campbell, David DeVries, Laura Emptage, Christopher Cookson, Marsha Kisilak, Juan Bueno, and Francisco Avila. Polarization Properties of Amyloid Beta in the Retina of the Eye as a Biomarker of Alzheimer's Disease. In *Optics in the Life Sciences*, page BM3A.4, Washington, D.C., apr 2015. OSA.

[29] Bernhard Baumann, Adelheid Woehrer, Gerda Ricken, Marco Augustin, Christian Mitter, Michael Pircher, Gabor G Kovacs, and Christoph K Hitzenberger. Visualization of neuritic plaques in Alzheimer's disease by polarization-sensitive optical coherence microscopy. *Scientific reports*, 7(March):43477, 2017.

[30] Max Born, Emil Wolf, A. B. Bhatia, P. C. Clemmow, D. Gabor, A. R. Stokes, A. M. Taylor, P. A. Wayman, and W. L. Wilcock. *Principles of Optics*. Cambridge University Press, Cambridge, 1999.

[31] Max Born and Emil Wolf. Chapter 15 Optics of crystals. In *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, pages 790–852. Cambridge University Press, 7th editio edition, 2000.

[32] Nate J. Kemp, Haitham N. Zaatari, Jesung Park, H. Grady Rylander III, and Thomas E. Milner. Form-biattenuance in fibrous tissues measured with polarization-

sensitive optical coherence tomography (PS-OCT). *Optics Express*, 13(12):4611, jun 2005.

[33] Jose Jorge Gil Perez and Razvigor Ossikovski. *Polarized light and the Mueller matrix approach*. CRC Press, New York, 2016.

[34] Christian Brosseau. *Fundamentals of polarized light : a statistical optics approach*. Wiley, 1998.

[35] Serge. Huard. *Polarization of light*. John Wiley, 1997.

[36] Zhang-Fan Xing. On the Deterministic and Non-deterministic Mueller Matrix. *Journal of Modern Optics*, 39(3):461–484, 1992.

[37] Shih-Yau Lu and Russell A. Chipman. Interpretation of Mueller matrices based on polar decomposition. *Journal of the Optical Society of America A*, 13(5):1106, may 1996.

[38] Dennis H. Goldstein. *Polarized light*. CRC Press,, Boca Raton, 2011.

[39] *Handbook of optics*. McGraw-Hill,, New York :, 2010.

[40] A Ambirajan and D C Look. Optimum angles for a Mueller matrix polarimeter. *Proceedings of the SPIE*, 2265(6):314–326, 1995.

[41] A Ambirajan and D C Look. Optimum angles for a Mueller matrix polarimeter. *Proceedings of the SPIE*, 2265(6):314–326, 1995.

[42] R.M.A. Azzam. Division-of-amplitude Photopolarimeter (DOAP) for the Simultaneous Measurement of All Four Stokes Parameters of Light. *Optica Acta: International Journal of Optics*, 29(5):685–689, may 1982.

[43] R. M. A. Azzam. Division-of-amplitude photopolarimeter based on conical diffraction from a metallic grating. *Applied Optics*, 31(19):3574, jul 1992.

[44] R. M. A. Azzam. Photopolarimetric measurement of the Mueller matrix by Fourier analysis of a single detected signal. *Optics Letters*, 2(6):148, jun 1978.

[45] Lixin Chin, Xiaojie Yang, Robert A McLaughlin, Peter B Noble, and David D Sampson. En face parametric imaging of tissue birefringence using polarization-sensitive optical coherence tomography. *Journal of Biomedical Optics*, 18(6):66005, 2013.

[46] Ying Yang, Asha Rupani, Pierre Bagnaninchi, Ian Wimpenny, and Alan Weightman. Study of optical properties and proteoglycan content of tendons by polarization sensitive optical coherence tomography. *Journal of biomedical optics*, 17(8):081417, 2012.

[47] M J Everett, K Schoenenberger, B W Colston, and L B Da Silva. Birefringence characterization of biological tissue by use of optical coherence tomography. *Optics letters*, 23(3):228–30, feb 1998.

[48] Chau-Chung Wu, Yih-Ming Wang, Long-Sheng Lu, Chia-Wei Sun, Chih-Wei Lu, Meng-Tsan Tsai, and C C Yang. Tissue birefringence of hypercholesterolemic rat liver measured with polarization-sensitive optical coherence tomography. *Journal of biomedical optics*, 12(6):064022, jan.

[49] Michael Pircher, Erich Götzinger, Oliver Findl, Stephan Michels, Wolfgang Geitzenauer, Christina Leydolt, Ursula Schmidt-Erfurth, and Christoph K. Hitzenberger. Human macula investigated in vivo with polarization-sensitive optical coherence tomography. *Investigative Ophthalmology and Visual Science*, 47(12):5487–5494, 2006.

[50] Stefan Zotter, Michael Pircher, Erich Götzinger, Teresa Torzicky, Hirofumi Yoshida, Futoshi Hirose, Stephan Holzer, Julia Kroisamer, Clemens Vass, Ursula Schmidt-Erfurth, and Christoph K. Hitzenberger. Measuring retinal nerve fiber layer birefringence, retardation, and thickness using wide-field, high-speed polarization sensitive spectral domain OCT. *Investigative Ophthalmology and Visual Science*, 54(1):72–84, 2013.

[51] Peijun Gong, Lixin Chin, Shaghayegh Es'haghian, Yih Miin Liew, Fiona M Wood, David D Sampson, and Robert A McLaughlin. Imaging of skin birefringence for human scar assessment using polarization-sensitive optical coherence tomography aided by vascular masking. *Journal of biomedical optics*, 19(12):126014, dec 2014.

[52] Wolfgang Drexler and James G. Fujimoto, editors. *Optical Coherence Tomography*. Springer International Publishing, Cham, 2015.

[53] Michael R. Hee, Eric A. Swanson, James G. Fujimoto, and David Huang. Polarization-sensitive low-coherence reflectometer for birefringence characterization and ranging. *Journal of the Optical Society of America B*, 9(6):903, jun 1992.

[54] Erich Gotzinger, Michael Pircher, Markus Sticker, Adolf F. Fercher, and Christoph K. Hitzenberger. Measurement and imaging of birefringent properties of the human

cornea with phase-resolved, polarization-sensitive optical coherence tomography. *Journal of Biomedical Optics*, 9(1):94, jan 2004.

[55] Erich Götzinger, Michael Pircher, and Christoph K. Hitzenberger. High speed spectral domain polarization sensitive optical coherence tomography of the human retina. *Optics Express*, 13(25):10217, dec 2005.

[56] S Makita, Y Yasuno, T Endo, M Itoh, and T Yatagai. Jones matrix imaging of biological samples using parallel-detecting polarization-sensitive Fourier domain optical coherence tomography. *Optical Review*, 12(2):146–148, 2005.

[57] Gang Yao and Lihong V. Wang. Two-dimensional depth-resolved Mueller matrix characterization of biological tissue by optical coherence tomography. *Optics Letters*, 24(8):537, apr 1999.

[58] Y. Yasuno, S. Makita, Y. Sutoh, M. Itoh, and T. Yatagai. Birefringence imaging of human skin by polarization-sensitive spectral interferometric optical coherence tomography. *Optics Letters*, 27(20):1803, oct 2002.

[59] Razvigor Ossikovski. Analysis of depolarizing Mueller matrices through a symmetric decomposition. *Journal of the Optical Society of America A*, 26(5):1109, may 2009.

[60] Clément Fallet, Angelo Pierangelo, Razvigor Ossikovski, Antonello De Martino, Clément Fallet, Angelo Pierangelo, and Antonello De Martino. Experimental validation of the symmetric decomposition of Mueller matrices. *Opt. Express*, 2832(2004):974–976, 2009.

[61] Razvigor Ossikovski. Differential matrix formalism for depolarizing anisotropic media. *Optics Letters*, 36(12):2330, jun 2011.

[62] José Jorge Gil and Eusebio Bernabeu. Depolarization and Polarization Indices of an Optical System. *Optica Acta: International Journal of Optics*, 33(2):185–189, feb 1986.

[63] Rafael Espinosa-Luna and Eusebio Bernabeu. On the Q(M) depolarization metric. *Optics Communications*, 277(2):256–258, 2007.

[64] Oriol Arteaga, Enric Garcia-Caurel, and Razvigor Ossikovski. Anisotropy coefficients of a Mueller matrix. *Journal of the Optical Society of America A*, 28(4):548, apr 2011.

[65] David DeVries, Melanie C W Campbell, Laura Emptage, Chris Cookson, Marsha Kisilak, Francisco J. Avila, Juan M Bueno, Rachel Redekop, and Matthew Wilson. Polarization properties of amyloid beta deposits in ex vivo human retinas from those with Alzheimer's disease differ from surrounding retina. In *Investigative Ophthalmology & Visual Science*, volume 56, pages 2385–2385. C.V. Mosby Co, jun 2015.

[66] E. Goetzinger, M. Pircher, B. Baumann, C. Hirn, C. Vass, and C. K. Hitzenberger. Retinal nerve fiber layer birefringence evaluated with polarization sensitive spectral domain OCT and scanning laser polarimetry: A comparison. *Journal of Biophotonics*, 1(2):129–139, 2008.

[67] Qienyuan Zhou and Robert W Knighton. Light scattering and form birefringence of parallel cylindrical arrays that represent cellular organelles of the retinal nerve fiber layer. *Applied optics*, 36(10):2273–85, 1997.

[68] Clifford R. Jack, Marilyn S. Albert, David S. Knopman, Guy M. McKhann, Reisa A. Sperling, Maria C. Carrillo, Bill Thies, and Creighton H. Phelps. Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3):257–262, may 2011.

[69] Beta-Amyloid (1-42), Ultra Pure, NaOH — rPeptide.

[70] B. Srinivasa Reddy and B. N. Chatterji. An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE Transactions on Image Processing*, 5(8):1266–1271, 1996.

[71] The MathWorks. MATLAB and Image Processing Toolbox Release 2016b, 2016.

[72] Julian Besag. On the Statistical Analisys of Dirty Pictures. *Journal of the Royal Statistical Society. Series B (Methodological)*, 43(3):259–302, 1986.

[73] M Karnoukian, S Faisan, Ch. Heinrich, A Lallement, and J Zallat. Segmentation of Mueller matrix images under non-uniform illumination. *Opt. Express*, 23(14):18218–18235, 2015.

[74] Christophe Collet, Jihad Zallat, and Yoshitate Takakura. Clustering of Mueller matrix images for skeletonized structure detection. *Optics express*, 12(7):1271–1280, 2004.

[75] Giorgos Sfikas, Christian Heinrich, Jihad Zallat, Christophoros Nikou, and Nikos Galatsanos. Recovery of polarimetric Stokes images by spatial mixture models. *Journal of the Optical Society of America A*, 28(3):465, 2011.

[76] Tony F. Chan and Luminita A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.

[77] Carl Zeiss. ZEN 2.3 lite.

[78] A. Elhaddaoui, E. Pigorsch, A. Delacourte, and S. Turrell. Competition of congo red and thioflavin S binding to amyloid sites in alzheimer's diseased tissue. *Biospectroscopy*, 1(5):351–356, 1995.

[79] Masahiro Yamanari, Masahiro Miura, Shuichi Makita, Toyohiko Yatagai, and Yoshiaki Yasuno. Phase retardation measurement of retinal nerve fiber layer by polarization-sensitive spectral-domain optical coherence tomography and scanning laser polarimetry. *Journal of Biomedical Optics*, 13(1):014013, 2008.

[80] H Kergoat, M J Kergoat, L Justino, H Chertkow, A Robillard, and H Bergman. An evaluation of the retinal nerve fiber layer thickness by scanning laser polarimetry in individuals with dementia of the Alzheimer type. *Acta ophthalmologica Scandinavica*, 79(2):187–91, apr 2001.

[81] H V Danesh-Meyer, H Birch, J Y-F Ku, S Carroll, and G Gamble. Reduction of optic nerve fibers in patients with Alzheimer disease identified by laser imaging. *Neurology*, 67(10):1852–4, nov 2006.

[82] Melanie C. Campbell, David Devries, Theodore Chow, Laura Emptage, Namrata Shah, Howard Dobson, Christopher Cookson, and Marsha Kisilak. Polarization Imaging in the Retina of a Dog Model of Alzheimer's Disease Enables Discrimination of Amyloid Deposits. In *Frontiers in Optics 2015*, page FTu5B.4, Washington, D.C., oct 2015. OSA.

[83] G. B. Benedek. Theory of Transparency of the Eye. *Applied Optics*, 10(3):459, mar 1971.

[84] Mireille Delaye and Annette Tardieu. Short-range order of crystallin proteins accounts for eye lens transparency. *Nature*, 302(5907):415–417, mar 1983.

[85] W. L. Bragg, A. B. Pippard, and IUCr. The form birefringence of macromolecules. *Acta Crystallographica*, 6(11):865–867, nov 1953.

[86] E W Taylor and W Cramer. Birefringence of protein solutions and biological systems. II. Studies on TMV, tropocollagen, and paramyosin. *Biophysical journal*, 3:143–54, mar 1963.

[87] E W Taylor and W Cramer. Birefringence of protein solutions and biological systems. I. *Biophysical journal*, 3:127–41, mar 1963.

[88] Melanie C W Campbell. Method and system for imaging amyloid beta in the retina of the eye in association with alzheimer's disease, may 2011.

[89] Jérôme Morio and François Goudail. Influence of the order of diattenuator, retarder, and polarizer in polar decomposition of Mueller matrices. *Optics Letters*, 29(19):2234, oct 2004.

[90] M. J. Kronenburg. A Method for Fast Diagonalization of a 2x2 or 3x3 Real Symmetric Matrix. 1:1–9, 2013.

[91] The MathWorks. MATLAB and Parallel Computing Toolbox Release 2018a, 2018.

[92] Nobuyuki Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, jan 1979.

[93] Stuart P. Lloyd. Least Squares Quantization in PCM. *IEEE Transactions on Information Theory*, 28(2):129–137, 1982.

[94] Geoffrey J. McLachlan and David Peel. *Finite mixture models*. Wiley, 2000.

[95] Keinosuke. Fukunaga. *Introduction to statistical pattern recognition*. Academic Press, 1990.

[96] Dongju Liu and Jian Yu. Otsu method and K-means. *Proceedings - 2009 9th International Conference on Hybrid Intelligent Systems, HIS 2009*, 1(2):344–349, 2009.

[97] Trevor Hastie, Robert Tibshirani, and J. H. (Jerome H.) Friedman. *The elements of statistical learning : data mining, inference, and prediction*. Springer-Verlag New York, New York.

[98] Arthur P. Dempster, Nan M. Laird, and Donald B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society, Series B.*, 39(1):1–38, 1977.

[99] Omer. Demirkaya, Musa Hakan. Asyali, and Prasanna. Sahoo. *Image processing with MATLAB : applications in medicine and biology*. CRC Press, 2009.

[100] John; Hammersley and Peter; Clifford. Markov fields on finite graphs and lattices. 1971.

[101] R. B. Potts and C. Domb. Some generalized order-disorder transformations. *Mathematical Proceedings of the Cambridge Philosophical Society*, 48(01):106, jan 1952.

[102] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. may 2015.

[103] Rachel; Redekop, Christopher; Cookson, David; DeVries, and Melanie C W; Campbell. Imaging hemozoin and amyloid beta in the retina of individuals with malaria using polarised light. In *Investigative Ophthalmology & Visual Science*, volume 58, pages 3366–3366. C.V. Mosby Co, jun 2017.

[104] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mane, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viegas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. mar 2016.

[105] Https://github.com/pytorch/pytorch. PyTorch.

APPENDICES

Appendix A

Cayley-Hamilton theorem for a 3×3 matrix

Cayley-Hamilton theory is an important theory in understanding a key step in the polar decomposition and also one of the steps that the accelerated implementation, shown in Chapter 3, can manipulate to boost the speed of computation. Here, a derivation of the Cayley-Hamilton theory applying to a 3×3 matrix $\mathbf{m}_{\Delta}$ in Section 3.10 is shown. This derivation follows [38].
For simplicity, the $\mathbf{m}_{\Delta}$ is represented by $\mathbf{m}$ in this section. The characteristic polynomial of $\mathbf{m}$ is

$$\begin{aligned}\det(\mathbf{m}-\lambda\mathbf{I}) =& \begin{vmatrix} m_{00}-\lambda & m_{01} & m_{02} \\ m_{10} & m_{11}-\lambda & m_{12} \\ m_{21} & m_{21} & m_{22}-\lambda \end{vmatrix} \\ =&\lambda^3-(m_{00}+m_{11}+m_{22})\lambda^2+(m_{00}m_{11}+m_{11}m_{22}+m_{00}m_{22} \\ &-m_{01}m_{10}-m_{02}m_{20}-m_{12}m_{21})\lambda-(m_{00}m_{11}m_{22}+m_{01}m_{12}m_{20} \\ &+m_{02}m_{21}m_{10}-m_{01}m_{10}m_{22}-m_{02}m_{20}m_{11}-m_{12}m_{21}m_{00})\end{aligned} \tag{A.1}$$

To simplify the above equation, one needs to use the properties of the determinant and the trace of $\mathbf{m}$. The determinant of $\mathbf{m}$ is exactly the zeroth order term in Eq.A.1, and it can be represented by the eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$ of $\mathbf{m}$ as

$$\begin{aligned}\det(\mathbf{m}) =& \prod eig(\mathbf{m}) \\ =&\lambda_1\lambda_2\lambda_3\end{aligned} \tag{A.2}$$

Second the trace of a matrix equals the sum of its eigenvalues. Hence, $\mathrm{tr}(\mathbf{m})$ can be written as

$$\begin{aligned}\mathrm{tr}(\mathbf{m}) =& m_{00} + m_{11} + m_{22} \\ =& \lambda_1 + \lambda_2 + \lambda_3\end{aligned} \tag{A.3}$$

So the coefficient of λ_2 in A.1 can be replaced by the eigenvalues.
Then, $(\mathrm{tr}(\mathbf{m}))^2$ and $\mathrm{tr}(\mathbf{m}^2)$ are

$$(\mathrm{tr}(\mathbf{m}))^2 = m_{00}^2 + m_{11}^2 + m_{22}^2 + 2(m_{00}m_{11} + m_{00}m_{22} + m_{11}m_{22}) \tag{A.4}$$

$$\mathrm{tr}(\mathbf{m}^2) = m_{00}^2 + m_{11}^2 + m_{22}^2 + 2(m_{01}m_{10} + m_{02}m_{20} + m_{12}m_{21}) \tag{A.5}$$

It is clear that the coefficient of λ in Eq.A.1 is just $\frac{1}{2}(\mathrm{tr}(\mathbf{m}))^2 - \mathrm{tr}(\mathbf{m}^2)$. Further, $(\mathrm{tr}(\mathbf{m}))^2$ and $\mathrm{tr}(\mathbf{m}^2)$ also can be written in terms of $(\lambda_1, \lambda_2, \lambda_3)$ as

$$\begin{aligned}(\mathrm{tr}(\mathbf{m}))^2 =& (\lambda_1 + \lambda_2 + \lambda_3)^2 \\ =& \lambda_1^2 + \lambda_2^2 + \lambda_3^2 + 2(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3) \qquad \text{(A.6)} \\ \mathrm{tr}(\mathbf{m}^2) =& \sum eig(\mathbf{m}^2) \\ =& \lambda_1^2 + \lambda_2^2 + \lambda_3^2 \qquad \text{(A.7)}\end{aligned}$$

Then the coefficient of λ in Eq.A.1 becomes $(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3)$.
Hence, plugging Eq.A.2, A.6, A.7 into Eq.A.1 will result in

$$\lambda^3 - (\lambda_1 + \lambda_2 + \lambda_3)\lambda^2 + (\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3)\lambda - \lambda_1\lambda_2\lambda_3 = 0 \tag{A.8}$$

The Cayley-Hamilton theory states that a square matrix satisfies its own characteristic polynomial. Hence, replacing λ by $\mathbf{m}$ in Eq.A.8, one can get

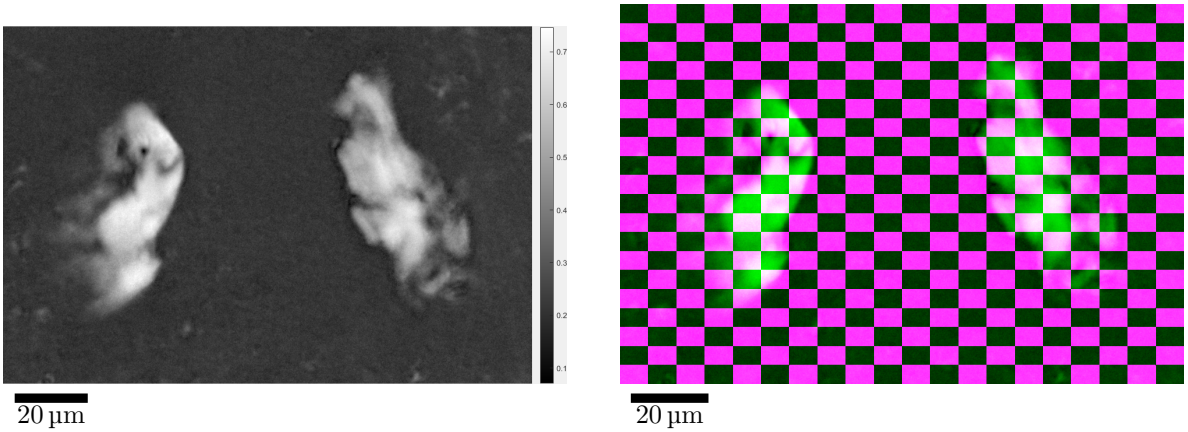
$$\mathbf{m} = [\mathbf{m}^2 + (\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3)\mathbf{I}]^{-1}[(\lambda_1 + \lambda_2 + \lambda_3)\mathbf{m}^2 + \lambda_1\lambda_2\lambda_3\mathbf{I}] \tag{A.9}$$

which is the same as Eq.3.10.

Appendix B

Active contour without edge

In Chapter 2, 3D images acquired by CLSM were segmented from the surrounding retina using the method called active contour without edge [76]. When applying to 3D images, MATLAB [71] offers a good function "activecontour" which was used to segment the deposits for creating thickness. However, the source code for this function is not accessible and it cannot be implemented on 2D multichannel images. To demonstrate the working principle of this method, a sample code has been written by following the original paper [76] and used in this Appendix. It also works on multichannel images but its accuracy requires further comparison with the MRF method in Chapter 4.
For simplicity, this example is given in 2D on one A_L image. Suppose the deposit in Fig.B.1.(a) needs to be segmented and a primary mask was randomly generated using the 20×20 checkerboard in Fig.B.1.(b).



(a) A_L (b) Initial checkerboard mask

Figure B.1: (a) Image of A_L and (b) initial checkerboard mask

In practice, this checkerboard can be more detailed (for example 100×100 instead of 20×20) to make sure part of the deposit overlap with the edges of checkerboard. The active contour method given by [76] aims to find a contour C to minimize the functional in Eq.B.1 of image $f(x)$. ($f(x)$ is expressed as the intensity of this image which is a function of spatial coordinates x)

$$C = arg\min_{C} \ \mu\text{Length}(C) + \nu\text{Area}(\text{inside}(C)) + \lambda_1 \int_{\text{inside}(C)} |f(x) - c_1|^2 dx + \lambda_2 \int_{\text{outside}(C)} |f(x) - c_2|^2 dx \tag{B.1}$$

where C is the contour and c_1, c_2 are regional averages inside and outside the contour. $\mu, \nu, \lambda_1, \lambda_2$ are predefined constants. The MATLAB function does not requires user input of these constants hence it may be that these constants have optimal settings in practice. The contour C is initialized as the edge of the checkerboard mask in Fig.B.1.(b). To simplify the minimization problem, this contour C is implicitly represented as the zero level of a level set $\varphi(x)$. This level set $\varphi(x)$ is defined as the euclidean distance of a point to the edge of the contour C. In MATLAB, this can be achieved by applying a binary distance transform both inside and outside the contour. The result of applying the distance transform on the checkerboard mask in Fig.B.1.(b) is a distance map shown in Fig.B.2.

Figure B.2: Distance map of the mask in Fig.B.1.(b)

The distance map inside the checkerboard mask is negative while the outside is positive. Then, the area term and length term in Eq.B.1 can be expressed in terms of the level set function $\varphi(x)$ as

$$\text{Area}(\text{inside}(C)) = \int H(\varphi(x))dx \tag{B.2}$$

$$\begin{aligned} \text{Length}(C) &= \int |\nabla H(\varphi(x))|dx \\ &= \int \delta(\varphi(x))|\nabla\varphi(x)|dx \end{aligned} \tag{B.3}$$

where $H(\varphi(x))$ is a step function which equals 1 if $\varphi(x)$ is larger than 0. Otherwise it equals 0. Its gradient is the Dirac delta function $\delta(\varphi)$.

Hence, Eq.B.1 becomes

$$C = arg\min_{C} \ \mu \int \delta(\varphi(x))|\nabla\varphi(x)|dx + \nu \int H(\varphi(x))dx + \lambda_1 \int |f(x) - c_1|^2 H(\varphi(x))dx + \lambda_2 \int |f(x) - c_2|^2 (1 - H(\varphi(x)))dx \tag{B.4}$$

where the regional averages c_1 and c_2 are

$$c_1 = \frac{\int f(x)H(\varphi(x))dx}{\int H(\varphi(x))dx} \tag{B.5}$$

$$c_2 = \frac{\int f(x)(1 - H(\varphi(x)))dx}{\int (1 - H(\varphi(x)))dx} \tag{B.6}$$

The minimization of Eq.B.4 can be achieved by applying the Euler-Lagrange equation to Eq.B.4. The result is an iterative method of

$$\frac{\partial\varphi}{\partial t} = \delta(\varphi)[\mu\nabla \cdot (\frac{\nabla\varphi(x)}{|\nabla\varphi(x)|} - \mu - \lambda_1(f(x) - c_1)^2 + \lambda_2(f(x) - c_2)^2] \tag{B.7}$$

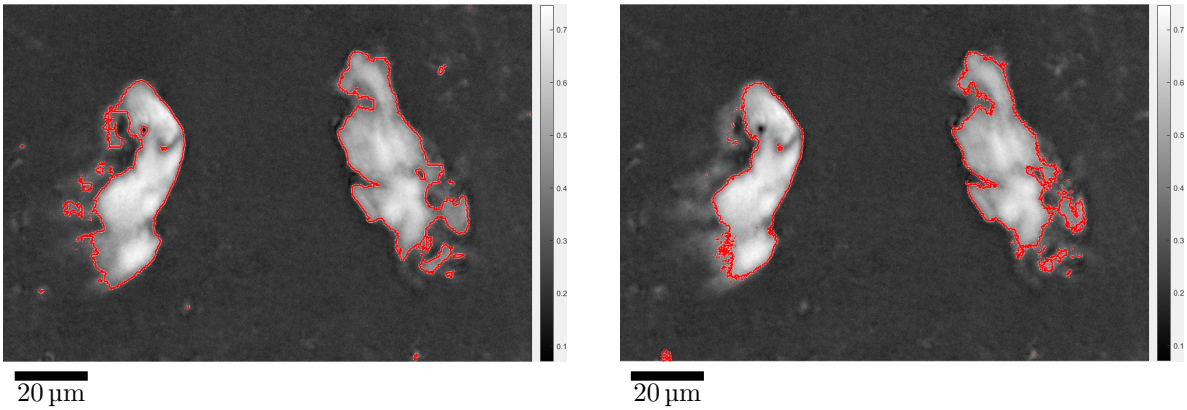
where t is the iteration step. Evolving the level set φ can be achieved by the finite difference method which gives

$$\varphi_{t+1} = \varphi_t + \frac{\partial\varphi}{\partial t} \tag{B.8}$$

After a number of iterations, the final level set is computed. The final mask is computed in Eq.B.9 as the region in the final level set $\varphi(x)$ that is smaller than 0 based on the definition of the level set defined in Fig.B.2.

$$\text{mask} = \varphi(x) < 0 \tag{B.9}$$

The result of applying the iterative method on the A_L image in Fig.B.1(a) is shown in Fig.B.3.(a). This method can also be adjusted to apply to multichannel images. The result of applying to multichannel image in Chapter 4 is shown in Fig.B.3.(b).



(a) A_L (b) Multichannel image

Figure B.3: (a) Final mask displayed on the A_L image after applying the active contour method to an A_L image. (b) Final mask displayed on the A_L image after applying the active contour method to the multichannel image introduced in Chapter 4.

As shown in Fig.B.3, the result of applying the method to the multichannel image is similar to the result of applying it to A_L image. The MATLAB function "activecontour" was applied to segmenting the 3D CLSM images in Chapter 2 and was very successful.